

**DEVELOPMENT OF A NOVEL TECHNIQUE FOR SKEW  
DETECTION, ESTIMATION AND CORRECTION OF  
SCANNED DOCUMENT**

Thesis submitted towards the partial fulfilment of the requirements for the  
award of degree of

**MASTER OF ENGINEERING  
IN  
WIRELESS COMMUNICATION**

Submitted By

**Mukul Khajuria**

**Roll No. 801363019**

Under the guidance of

**Dr. Vinay Kumar**

**Assistant Professor, ECED**

**Thapar University, Patiala**



**ELECTRONICS AND COMMUNICATION ENGINEERING  
DEPARTMENT  
THAPAR UNIVERSITY**

**(Established under the section 3 of UGC Act, 1956)**

**PATIALA – 147004, PUNJAB, INDIA**

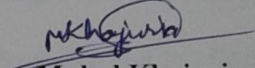
**July 2015**

## DECLARATION

I, **Mukul Khajuria**, hereby declare that the work, which is being presented in the thesis entitled "**Development of a Novel Technique for Skew Detection, Estimation and Correction for Scanned Document**" by me in partial fulfillment of the requirements for the award of degree of Master of Engineering in Wireless Communication from Thapar University, Patiala, is an authentic record of my own work carried out under the supervision of **Dr. Vinay Kumar**, Assistant Professor, Electronics and Communication Engineering Department.

The matter presented in this thesis has not been submitted in any other University/Institute for the award of any other degree.

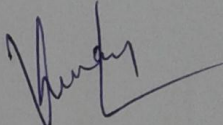
Date: 13-07-2015

  
**Mukul Khajuria**

Roll no: 801363019

This is to certify that the above statement made by the student is correct to the best of my knowledge and belief.

Date: 13-07-2015



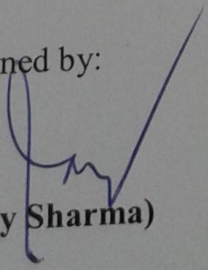
**Dr. Vinay Kumar**

Assistant Professor

ECED

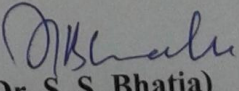
Thapar University, Patiala

Countersigned by:

  
**(Dr. Sanjay Sharma)**

Professor and Head, ECED

Thapar University, Patiala

  
**(Dr. S. S. Bhatia)**

Dean of Academic Affairs

Thapar University, Patiala

## ACKNOWLEDGEMENT

---

I would never been able to finish my thesis without the supervision of faculty of my department, help from friends, and provision from my family.

I would like to utter my deepest thankfulness to my guide, **Dr. Vinay Kumar**, Assistant Professor ECED, Thapar University, Patiala, for his excellent guidance, caring, patience, and providing me with an excellent atmosphere for doing research.

I would like to thank **Dr. Sanjay Sharma**, Professor and Head, ECED, Thapar University, Patiala, for granting us with sufficient facilities and infrastructure to carry our work.

I would like to acknowledge **Dr. Hemdutt Joshi**, Assistant Professor and Program Coordinator (WC), ECED, Thapar University, Patiala, for being a source of inspiration.

I would like to thank **Dr. Amit Kumar Kohli**, Associate Professor and PG Coordinator, ECED, Thapar University, Patiala, for his motivation and support.

I would also like to thank **Bhupinder Singh, Abhishek Bansal** and **Sagar Goyal**, who as good friends were always willing to help and give their best suggestions.

Finally, I would thank my parents and sister for their support, inspiration and motivation.

**Place: TU, Patiala**

**Date:**

**Mukul Khajuria**

**Roll No 801363019**

## ABSTRACT

---

We propose a novel technique for automatic skew detection, estimation, and correction. These defects are introduced in a document while scanning or capturing its picture. The process of automatic detection, estimation, and correction is involved in many interconnected operations.

In the present work, we are able to remove any skewness introduced in a document with or without images, having bar codes, bar graphs, pie charts, logos, tags, tables, etc.

The estimation of skew angle is calculated with help of probabilistic distribution of white and black pixels. Once the estimation is performed, the correction is carried on original document.

Our technique is applicable for wide range of documents, which was not possible with state-of-the-art techniques available.

**Keywords:** *Skew, skew angle, skew estimation, skew correction, OCR (optical character recognition), image rotation, document image processing, pre-processing, scanned documents.*

# TABLE OF CONTENTS

---

---

<b>DECLARATION</b>	<b>i</b>
<b>ACKNOWLEDGEMENT</b>	<b>ii</b>
<b>ABSTRACT</b>	<b>iii</b>
<b>TABLE OF CONTENTS</b>	<b>iv</b>
<b>LIST OF ABBREVIATIONS</b>	<b>vii</b>
<b>LIST OF FIGURES</b>	<b>viii</b>
<b>1. INTRODUCTION</b>	<b>1-17</b>
1.1. Image Processing	1
1.2. Steps in Digital Image Processing	2
1.2.1. Image Acquisition	3
1.2.2. Image Enhancement	4
1.2.3. Image Restoration	5
1.2.4. Colour Image Processing	5
1.2.5. Wavelets and Multi-Resolution Processing	5
1.2.6. Image Compression	6
1.2.7. Morphological Processing	7
1.2.8. Image Segmentation	8
1.2.9. Representation and Description	8
1.2.10. Image Recognition	9
1.3. E-Content	10
1.3.1. E-Book	10

1.3.2. E-Reader	10
1.4. Optical Character Recognition	11
1.4.1. Types of OCR	12
1.5. Advantages of OCR	14
1.6. Skew Estimation and Correction	15
<b>2. LITERATURE REVIEW</b>	<b>18-31</b>
2.1. Skew Detection and Estimation Approach based on Hough Transform	18
2.2. Skew Detection and Estimation Approach based on Projection Profile	21
2.3. Skew Detection and Estimation Approach based on Nearest Neighbour	25
2.4. Skew Detection and Estimation based on Log-Polar Transform	26
2.5. Skew Estimation Approach based on Interline Cross Correlation Clustering	27
2.6. Skew Estimation based on Fused Techniques	27
2.7. Skew Estimation based on PPA	28
2.8. Skew Estimation based on Image Dilation	29
2.9. Skew Correction Methods	29
2.9.1. Direct Methods	29
2.9.2. Indirect Methods	30
2.9.3. Contour Oriented Methods	30
2.10. Skew Correction using Cyclostyle Matching	31
<b>3. ANALYSIS AND METHODOLOGY</b>	<b>32-40</b>
3.1. Introduction	32
3.2. Implementation	33
3.3. Features of the Technique	40

<b>4. RESULTS AND DISCUSSIONS</b>	<b>41-53</b>
<b>5. CONCLUDING REMARKS AND FUTURE SCOPE</b>	<b>54</b>
5.1. Conclusion	54
5.2. Future Work	54
<b>REFERENCES</b>	<b>55-58</b>

## LIST OF ABBREVIATIONS

---

RGB	Red Green Blue Colour Space
YCbCr	Luminance Blue Chrominance Red Chrominance
HSV	Hue Saturation Value Colour Space
CWT	Continuous Wavelet Transform
DWT	Discrete Wavelet Transform
DPCM	Differential Pulse Code Modulation
LZW	Lempel-Ziv-Welch Image Compression Encoding
TIFF	Tagged Image File Format
GIF	Graphics Interchange Format
LCD	Liquid Crystal Display
OCR	Optical Character Recognition
OWR	Optical Word Recognition
ICR	Intelligent Character Recognition
IWR	Intelligent Word Recognition
PDF	Portable Document Format
SVM	Support Vector Machine
PPA	Piece-Wise Painting Algorithm

## LIST OF FIGURES

FIGURE NO.	TITLE OF THE FIGURE	PAGE NO.
Fig. 1.1.	Fundamental steps of digital image processing	1
Fig. 1.2.	A sensor to generate 2-D image	3
Fig. 1.3.	Seismic wavelet	6
Fig. 1.4.	Skewed text image	15
Fig. 1.5.	De-skewed text image	16
Fig. 1.6.	Original image in correct alignment	16
Fig. 1.7.	Original image skewed by 7 degrees	17
Fig. 2.1.	Horizontal projection profiles of de-skewed image	22
Fig. 2.2.	Vertical projection profiles of de-skewed image	22
Fig. 2.3.	Horizontal projection profiles of skewed image	23
Fig. 2.4.	Vertical projection profiles of skewed image	23
Fig. 3.1.	Skewed image	34
Fig. 3.2.	Skewed image rotated by 2 degrees	35
Fig. 3.3.	Complemented skewed image	36
Fig. 3.4.	Text image within boundary	37
Fig. 3.5.	Cropped image	38
Fig. 3.6.	De-skewed image.	38
Fig. 3.7.	Resultant de-skewed image	39
Fig. 4.1.1.	Skewed document image with text and logos	41
Fig. 4.1.2.	De-skewed document image with text and logos	42
Fig. 4.2.1.	Skewed document image with logos, text and tables	43
Fig. 4.2.2.	De-skewed document image with logos, text and tables	44
Fig. 4.3.1.	Skewed document image of characters	45
Fig. 4.3.2.	De-skewed document image of characters	45
Fig. 4.4.1.	Skewed document having only image (with boundary)	46
Fig. 4.4.2.	De-skewed document having only image (with boundary)	47
Fig. 4.5.1.	Skewed handwritten text document in sanskrit language	48
Fig. 4.5.2.	De-skewed handwritten text document in sanskrit language	49
Fig. 4.6.1.	Skewed image of bar code	50

Fig. 4.6.2.	De-skewed image of bar code	51
Fig. 4.7.1.	Skewed image of a bar graph	52
Fig. 4.7.2.	De-skewed image of a bar graph	52
Fig. 4.8.1.	Skewed image of a distribution	53
Fig. 4.8.2.	De-skewed image of a distribution	53

# INTRODUCTION

---

## 1.1 Image Processing

Image processing is used for performing operations on the original image in order to obtain an image with improved quality or to gather out some relevant information from it. It is a classification of dispensation of signal in two-dimension, in which image is considered as input, like a frame of video or photograph, and its output can be an image or characters of it that are related with that image. Generally, image processing systems treat images as a signal in two-dimension while various methods of signal processing are being applied to them. It is one of the fast growing technologies that are present today, with its wide range of applications.

In general, image processing comprises of the following three steps:

- 1) Capturing image with the help of digital photography or by using optical scanner.
- 2) Manipulating and analysing the image which includes compression of the data within the image and enhancing the quality of image.
- 3) Analyse the output in which results can be altered image.

Image processing is used for better visualization purpose; to detect and analyse the things, as such which are not visible. Further sharpening and restoration of image is applied to create an enhanced image. Seeking the image of interest, image retrieval can be done. Further we can apply image processing for the measurement of various objects and patterns in an image and also for distinguishing the objects in an image for better image recognition.

The use of algorithms used in computer, for performing the processing on images that are available in digital form is known as digital image processing. Digital image processing is classified as a subclass of digital signal processing. The technique of processing a digital image has many advantages over the processing techniques used for analog images. We apply an extensive variety of algorithms to the data that is input, in digital image processing and further avoids the problem as that of noise and distortion of signal that is being generated during the signal processing. Since the images are defined over two-dimensions, the processing techniques used in processing of digital images are modelled in the form of multidimensional systems.

The images in a digital imaging system are formed by a countable number of elements and each of them having a specific value and location; and are stated as image elements, elements of picture and pixels. For representing the elements of a digital image, the most widely used term is pixel. Today, in some way or other, almost every area of technical endeavour is impacted by digital image processing.

## 1.2 Steps in Digital Image Processing

It is convenient to distribute the material enclosed further into two broad categories that are defined as: approaches having images as their inputs but the extracted attributes of the images are the outputs; and the methods whose input as well as output both are images. This structure is summarized and illustrated in Fig. 1.1 below.

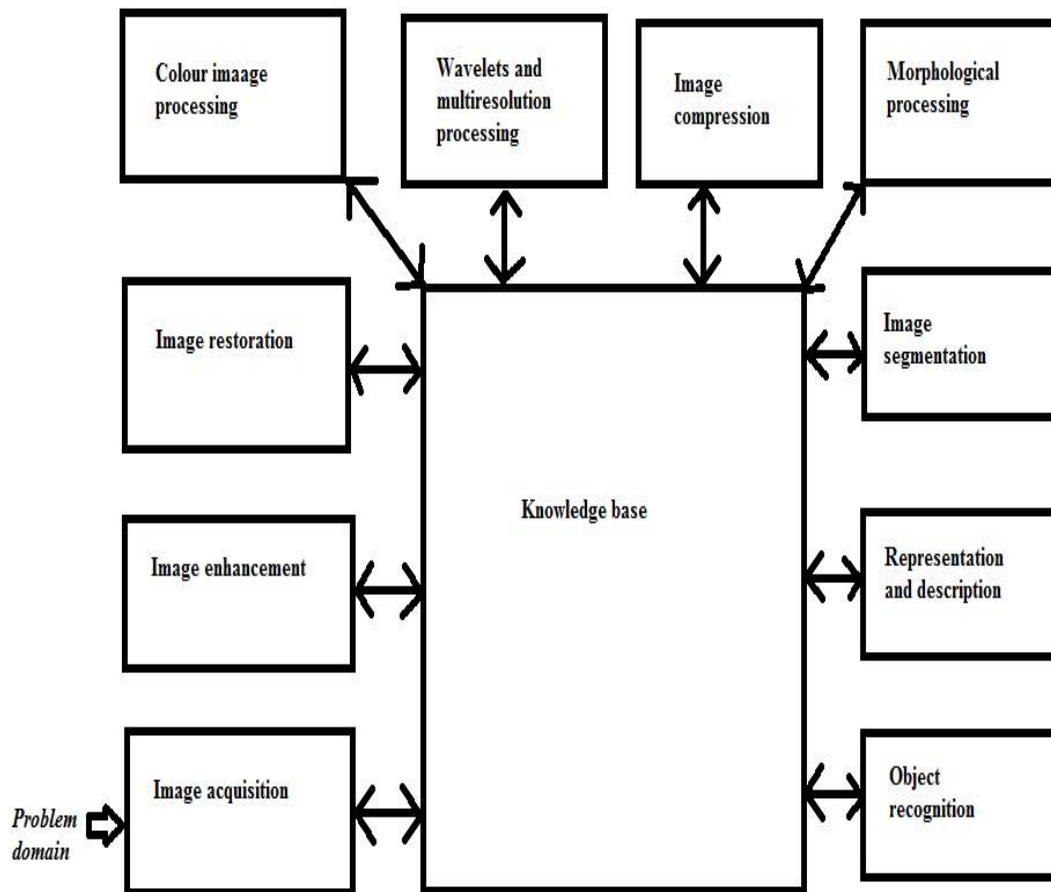


Figure 1.1. Fundamental steps of digital image processing [1].

The implication of this summarized figure does not reveal that each process in it is applicable directly or indirectly to an image. Basically the purpose is to deliver the

concept of all approaches and techniques that are analysed on pictures with altered perspective.

### 1.2.1 Image Acquisition

In image acquisition, the image is previously in digital form. The first process shown in Fig.1 is image acquisition. The stage of image acquisition generally involves pre-processing, such as scaling. The acquisition of an image can be done using single sensor, sensor strips and also using sensor arrays. In image acquisition with the use of single sensor, the components of a single sensor are as shown in Figure 1.2.

Photodiode is the most common and familiar sensor of this type that is made of materials of silicon whose waveform of voltage at the output is proportional to light. The selectivity is improved by using a filter ahead of a sensor.

Consider an example as in front of light sensor, apply a filter (that passes only blue light) favours light in blue band of colour spectrum. So, its result is output of the sensor will be stronger for blue light than as from the other constituents of the visible spectrum. If a 2-D image has to be generated using a single sensor, then a relative displacement should be there in both ordinate and abscissa direction, between the part to be imaged and sensor. Figure 1.2 displays a sensor combination that is implemented for scanning a document with high precision, in which the negative of a film is fixed up on a drum and the rotation of the drum offers the one dimensional displacement.

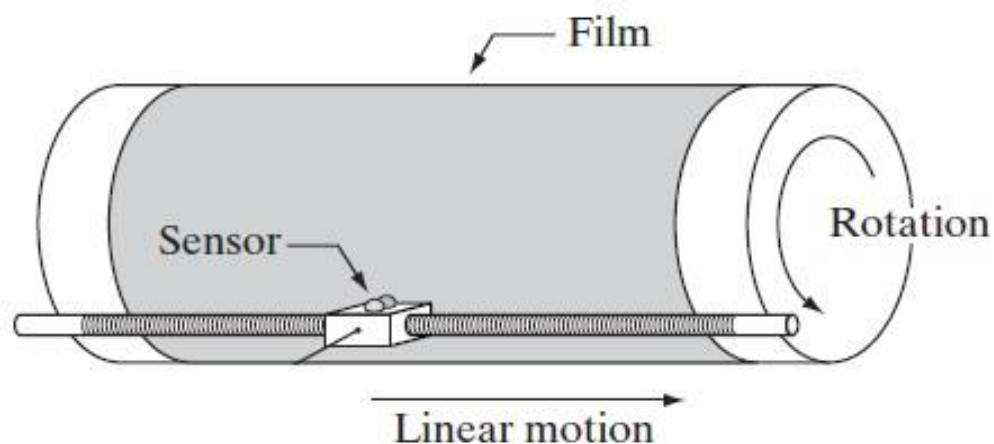


Figure 1.2. A sensor to generate 2-D image [1]

Here the sensor is mounted on lead screw, so that it should be able to provide motion in its perpendicular direction. This method of image acquisition is reasonable but is more

time consuming way for obtaining the images with more resolution, as mechanical motion can easily be controlled with high precision. There are some other similar mechanical arrangements that use flatbed scanners, in which the sensor moves in two linear directions and are known as microdensitometers.

### **1.2.2 Image Enhancement**

Image enhancement is one of the modest fields in digital image processing. The basic notion behind the methods used for enhancement of the image is to extract the details that are less visible in an image, or to lay emphasis on some features of interest of an image. An example of enhancement of image is as contrast within the image is increased for its better perception. Another example of image enhancement is the Fourier transformation. The main motive while enhancing an image is to process the image to obtain results that are additionally enhanced and more appropriate than that of previous image that was considered previously, for different applications.

For example, a method used for the enhancement of images of X-rays might not be considered as the best methodology to be used for the enhancement of images of Mars being sent by a space probe.

Enhancement of image is considered amongst the most visually appealing and interesting areas of processing of the image, regardless of the method used. There are two broad categories in which image enhancement approaches are divided as spatial and frequency domain. The spatial domain processing techniques relies on image plane, and methods here are on the basis of manipulation of pixels value in an image. The techniques that are used for processing in frequency domain rely on modification of the Fourier transform. As such there is not any universal theory available for image enhancement. While processing an image for interpreting it visually, the working of a particular method is judged. The visual recognition and estimation of quality of an image is actually a subjective process.

So, a good image can be defined as an elusive standard for comparing the performance of algorithm. In case when the problem arises in one of the image to be processed for perception using machine, the estimation task becomes easier to some extent. In case while handling the applications related to the character recognition, and without considering the supplementary issues as the requirements of computation, the image processing technique described to be the best, would be the one that yields the best results in machine recognition.

### **1.2.3 Image Restoration**

The image restoration deals with the improvement in form of an image. Unlike image enhancement, the restoration process in an image is objective; in regard to that the techniques used for restoration are based on probabilistic or mathematical prototypes used in process of degradation of image.

### **1.2.4 Colour Image Processing**

The area of colour image processing is presently gaining significance because of increasingly use of images available in digital form over the internet. A colour image is an image in digital form which comprises of the information about the colours for every value of pixel. For results to be accepted visually in colour images, it is mandatory to give three samples for each value of pixel, that are further revealed as some coordinates in colour space. The commonly used colour space for the computer displays today are RGB colour space, but other available colour spaces such as YCbCr or HSV can be used in other contexts. Three values per pixel are present in a colour image that is used in colour image processing has and the luminance and chrominance of light are measured by using those three values. The actual information which is being kept as data in digital image is the information about the brightness present in each band of spectrum.

### **1.2.5 Wavelets and Multi-Resolution Processing**

A wavelet is defined as a wave-like oscillation having amplitude that begins at zero, then increases, and decreases back to zero. It can be pictured as a small oscillation as that in case of a heart monitor or seismograph. Wavelets are shaped to ensure some specific properties by which they can be made beneficial for processing of signal. One can combine the wavelets by considering a technique that involves the following operation of “reverse”, “shift”, “multiply” and “integrate” together known as convolution, by considering some parts of the signal that is known previously to take out the data from a signal that is not known to us previously. A wavelet is a well-defined function that can be applied for dividing a continuous-time signal or a given function in various other scaled components. Usually a range of frequency is allocated for every component of that scale and further it can be studied with a precise resolution matching to its scale. The representation of a function by wavelets is called as wavelet transform.

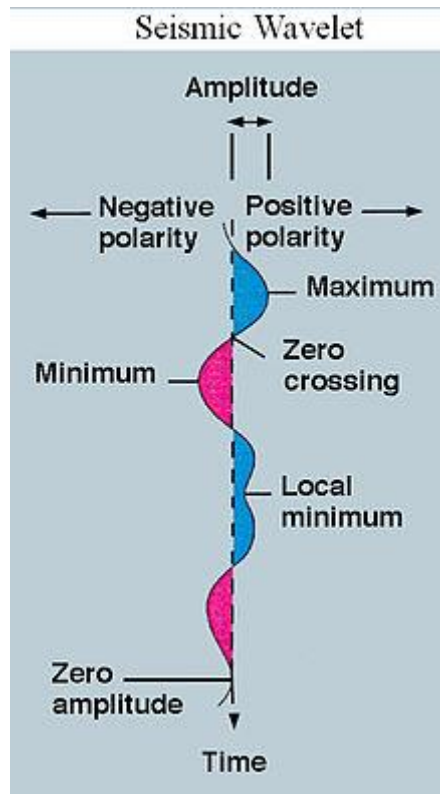


Figure 1.3. Seismic wavelet [2]

The wavelets are the scaled and translated copies of a fast-decaying or finite-length oscillating waveform. For the representation of functions having sharp peaks and discontinuities, the wavelet transform is beneficial and also for reconstructing and deconstructing the non-periodic, finite, and/or non-stationary signals accurately. The wavelet transforms are classified into two categories namely continuous wavelet transform (CWT'S) and discrete wavelet transform (DWT'S). It is noticed here that both CWT and DWT are transforms in continuous-time. The continuous time transform represents the continuous-time signals. DWT considers a precise subset of values of scale and translation or grid representation whereas CWT works on each feasible scale and translation.

### 1.2.6 Image Compression

Image compression can be defined as the encoding of data (or image) into another image which is represented by less number of bits so that data can be stored using less memory and transmitted efficiently. The inverse process of encoding is called decoding. The hardware and software system that can encode data are called encoders, and that decode the data are called decoders. They both combine and form a codec. For lossless image

compression, following methods can be used namely DPCM and predictive encoding, run-length encoding, entropy encoding, chain codes and also adaptive dictionary algorithms as that of LZW used in TIFF and GIF.

In case of lossy image compression, the methods used are reduction in colour space, chroma sub-sampling, transform coding, and fractal compression. The main goal of image compression is the improved quality of image at a given compression rate (bit-rate). A reduction in quality that is achieved by the manipulation of file or bit-stream is termed as scalability. It can also be termed as embedded bit-stream or progressive coding and can be seen in lossless codecs, mostly in that of coarse-to-fine scanning of pixels. It is beneficial for seeing images before downloading them using a browser or in an outcome to provide adjustable quality access.

There are various variants of scalability and are quality progressive, component progressive and resolution progressive. In quality progressive, the bit-stream refines the reconstructed image. In component progressive, image is first encoded into grey, then into a colour image. In resolution progressive, it is first encoded into an image with low resolution, then encode the difference to higher resolution.

### **1.2.7 Morphological Processing**

Morphology can be defined as set of operations used in processing of images that further process the images based on their shapes. The process used in morphological processing of an image; to an input image a structuring element is applied, and creates an output image with same size. In these processes, cost of each pixel present in the output image is actually based on the matching with the corresponding pixel present in the input image and its neighbouring pixels. By selecting a particular shape and size of the neighbourhood, an operation based on morphology can be constructed that is sensitive to particular shapes that are present in the input image.

The two main morphological operations available are erosion and dilation. Erosion operation removes pixels from an image which are present on the boundaries of objects while the operation of dilation adds the pixels to the boundaries of objects in an image. The pixels that are removed/added from/within the objects present in an image totally rely on the shape and size of the elements used for structuring is used in the processing of the image. In the morphological erosion and dilation operations, the given pixel's state in output image can be resolved easily by applying a set of instruction to the corresponding pixel and its neighbours that are present in the input image.

The set of rules used for the processing of pixels describes the operation as an erosion or dilation. In erosion, the output pixel value is the minimum value among all the pixels in the neighbourhood of input pixel. In case of a binary image, if any of the pixel value is set to be zero, then the output pixel value is set to zero. In dilation, the output pixel value is the maximum value among all the pixel values in the neighbourhood of input pixel. If any of the pixel value in a binary image is set to one, then the output pixel is set to one. The pixels that are beyond the border of image are assigned the minimum value yielded by the data type. In binary images, these pixels are supposed to be set to zero value.

### **1.2.8 Image Segmentation**

Image segmentation is the technique of division of digital image further into multiple segments. The main aim of image segmentation is to simplify the representation of an image into another image that is easier to analyse and more meaningful. Segmentation of images is done for finding the location of boundaries and objects in images. The output (or results) of image segmentation is a set of image segments that collectively cover the entire area in image. Every single pixel in a particular area is comparable with regard to various properties as that of intensity, colour, or texture or some other characteristics. Some applications of segmentation of image used in real-world are medical imaging, video surveillance, object detection, machine vision, recognition tasks, content-based image retrieval, and traffic control systems. The medical imaging further includes the diagnosis and study of anatomical structure, virtual surgery simulations, measure tissue volumes, surgery planning, and intra-surgery navigation. Object detection includes face detection, pedestrian detection, locating objects in images taken from satellite. Recognition tasks include fingerprint recognition, face recognition and iris recognition.

There are various methods for image segmentation and few of them are threshold method, compression based method, edge detection method, clustering method, histogram-based methods, region-growing methods, dual clustering methods, partial differential equation-based method and many others. The threshold method is simplest method of image segmentation in which a threshold value is adjusted for converting a greyscale image into a binary image.

### **1.2.9 Representation and Description**

Representation of an image is usually raw pixel data which constitutes either the boundary of a region or all the points in a region itself. It follows the output of image

segmentation stage. Converting the available data to a form that is suitable for further processing is necessary in either case. While representing an image, it has to be seen that data should be signified as a boundary or as a complete region.

When the main focus is on the external shape and characteristics, the boundary representation should be appropriate and when the focus is mainly on internal properties as that of skeletal shape and texture, then regional representation is appropriate. In few applications, these representations complement each other. Selecting a representation accordingly is only a solution for the transformation of raw data into a suitable form for later processing. Description (or feature selection), deals with the extraction of the attributes that leads to some information of interest and are very basic for differentiating one class of objects from the another class.

### **1.2.10 Image Recognition**

Image recognition is the process of detection and identification of a feature or an object in a digital image or video. It is also known as computer vision. It is used in the systems that are used in factories for automation purposes, in security surveillance systems, and also in toll booth monitoring. Basically the image recognition algorithm includes the pattern and gradient matching, optical character recognition, scene change detection and face recognition. Pattern and gradient matching is not limited to few types of images as that of facial and optical character recognition which can be used only for faces and text recognition respectively.

Pattern recognition is able to recognize any pattern of pixels. It can easily detect objects like cars cycles, specific images like the logo of a specific firm, or even objects arranged in different poses as that of a cartoon character. Face recognition lays emphasises on detection of faces and differentiates one face from another. For example, while uploading a photo on a social networking site, it prompts for tagging face in the picture with a particular friend name. This utilises the face recognition technology.

The main aim of optical character recognition technology is to detect the text within a document and also for reading it. While depositing check in the bank by taking its picture, the bank uses OCR for reading the routing and account numbers on that check so that funds should be transferred correctly from the correct account.

### **1.3 E- Content**

E-content means content in the electronic form is actually the digital content that can be transmitted over any computer network, for example the content transmission over the internet. Normally if we see the growth of technology applications in field of education, we are moving towards an era of virtual reality where there will be no distance between the teacher and the taught. Such virtual realities can be made possible by generating and developing good e-Contents which will be accessible by all. Basically e-contents are the package that satisfies the conditions like cost effectiveness, minimization of the distance for easy access and adaptability to local conditions. E-content should also be user friendly.

E-content is information in digital form that is sent over electronic devices that are based on network. It can be defined as the digital text and digital images which are designed for the display on web pages. Electronic-content is a combination of text, images, audio, video and also the animation with visual effects. Any type of digital content that facilitates the learning process and learning outcome is termed as e-content.

The acquisition of e-contents can be done via four different channels:

- 1) Purchasing of the materials.
- 2) Using the freely available content on the internet.
- 3) Self-production of the material.
- 4) Exchange of existing material in a network with other materials available on the other networks.

#### **1.3.1 E-Book**

An electronic book is a book published in digital form that includes the texts, images or both, which is readable on fixed or other handheld electronic devices. Many times e-book can be defined as an electronic version of a printed book. Many e-books are there without any printed equivalent of it. Some e-books which are produced and sold commercially are usually intended to be read on dedicated electronic readers (e-readers). Any sophisticated electronic device featuring a viewing screen, including computers, smartphones, tablets and many more handheld devices can also be used to read e-books.

#### **1.3.2 E-Reader**

An e-reader (e-book reader) is a mobile device that is intended particularly for the purpose of reading digital e-books. A device that displays text on a screen, acts as an e-

reader, but specialized e-readers are made considering the optimized designs, allowing proper portability, readability (brightness in sunlight), and battery life of the device for this purpose. A single electronic reader has the capacity of holding the hundreds of e-books with no added measurable mass to it.

E-readers are low-power, portable and high-resolution devices that are precisely designed for the display of digital versions of material written from the magazines, books, newspapers, and other printed sources. Various other e-readers are also available which gives access to the electronic documents like websites, blogs etc. The term e-reader refers here to a device designed specifically for presenting text. Rather than an LCD screen, many e-readers offer black and white text presented on so called e-ink and they also allow users to store a collection of titles in a single device. E-readers display the electronic versions of text, using a display technology which is particularly designed to simulate the printed paper. The resolution offered by e-reader is similar to the resolution as in newsprint and comparative to an LCD screen it eliminates the glare and also reduces strain on eyes.

Text document can be loaded through wireless or wired connections onto e-readers, from a referred website, or from an online store. Students are increasing the usage of e-readers nowadays for reading e-textbooks. The publishers of textbooks are laying emphasis on producing books in e-book formats.

E-readers and e-books are changing the economic trends for sellers and buyers of text based intellectual property, including the study materials. However electronic texts are less expensive yet, but there are presently limited options available for reselling them. The deviation of trend from textbooks to digital textbooks would have widened the implications for traditional bookstore and also for libraries, which stores the reserve materials. Many of those text documents that are not a copyright are available in electronic versions of document free of cost and hence they also offer a substantial savings to students when a particular edition is not required.

Advanced features such as hyperlinking, saved views, cross-linking, annotation and interactive quizzes for individual study. A large collection of books, notes and articles can be stored within a single e-reader. So, students need not to carry a bag full of textbooks and writing pads with them. Apart from these advantages of e-readers, some of them also provide text to speech conversion.

## **1.4 Optical Character Recognition**

Optical character recognition is the electronic or mechanical conversion of the images, a handwritten text document or printed text document; into text document that is encoded by machine. OCR is used as a type of data entry from printed paper data records, whether invoices, documents for passport, statements of bank, business cards, receipts generated by computer, static-data printouts, mail, or any other appropriate documentation. It is a technique used for digitizing the printed texts as that for storing it more compactly, edited by electronic means, examined, displayed on-line, and used in machine processes as in case of text skewing of skewed images, text-to-speech, machine translation etc.

OCR is a major field of research in computer vision, artificial intelligence and pattern recognition. Earlier variants of OCR were used to be accomplished with images of each character, and worked on only single font at a time. The innovative OCR systems have a high degree of accuracy for recognition of almost all fonts. There are specific OCR's available that efficiently reproduce the output that is formatted and nearly approximates the original page that includes the columns, non-textual components and images.

### **1.4.1 Types of OCR**

OCR is an offline process, in which document in static form is analysed. The concept analysing the movement of handwriting is used as an input for recognition of handwriting. This technique captures the motion, for example the order which draws the segments, the direction, and pattern of putting the pen down and lifting it, despite utilizing the shapes of glyphs and words. By using extra data available, end-to-end process can become more accurate. The technology based on OCR is also known as real-time character recognition, online character recognition, intelligent character recognition and dynamic character recognition.

Optical character recognition can be classified as below:

- 1) Optical Character Recognition (OCR) targets one character or glyph at a time in typewritten text.
- 2) Optical Word Recognition (OWR) targets one word at a time in typewritten text.
- 3) Intelligent Character Recognition (ICR) aims at the text written by hand and cursive text, one character at a time, frequently relating the machine learning to it.
- 4) Intelligent Word Recognition (IWR) aims the text written by hand and cursive text, one word or glyph at a time. It seems very beneficial for the languages in which glyphs are not separated in cursive script.

The procedure in optical character recognition includes the following four major steps as discussed below:

- 1) Pre-processing
- 2) Character recognition
- 3) Post-processing

- **Pre-processing**

The software of optical character recognition pre-processes the images to enhance their probability of effective recognition. These methods include-

- a) De-skewing of the text document which means while scanning a document if the text document was not properly aligned, then it is required to be skewed few degrees in counter clockwise direction or clockwise direction to sort lines of the text document perfectly horizontal or vertical.
- b) Removal of the positive and negative spots, and smoothing the edges of the text document.
- c) Binarization of the text document image which implies the transformation of an image from RGB to an image having only two colours that is black and white (or say a "binary image" because only two colours are present in it). It is mandatory for the character recognition algorithm, in some cases; and in other cases, the performance of algorithm is enhanced on the original image.
- d) Removal of the line scrubs up the non-glyph boxes and lines.
- e) Zoning or layout analysis recognises the paragraphs, columns, etc. as separate blocks. This is very significant in multi-column layouts.
- f) Line and word detection creates the baseline for the shapes of words and characters, and separate words if essential.
- g) Character segmentation or isolation is used for per-character OCR, multiple characters that are connected because of the image artefacts have to be separated and single characters, broken into multiple pieces because of artefacts have to be connected.
- h) In multilingual documents, script can vary at level of words. Hence, before invoking the OCR for handling specific scripts, identification and recognition of the script is necessary.

- **Character recognition**

There are two types of algorithms for OCR that produces a list of characters with specific ranks. In feature extraction, the decomposition of glyphs into "features" like lines, line direction, line intersections and closed loops. These can be associated with an abstract vector-like depiction of character, which is further reduced to one or more glyph prototypes. The normal methods of detecting the features in computer vision are inapplicable to this form of OCR that is generally observed in "intelligent handwriting recognition".

In matrix matching (or say "pattern matching" or "image correlation") the comparison of an image to a stored glyph is done on a pixel-by-pixel basis. This depends on correctly isolating the input glyph from all other images, and also on the stored glyph being in an analogous font and at the same scale. The pattern matching method gives best results only in case of the typewritten text and as the new fonts are encountered, it does not work well. Earlier, the physical photocell-based OCR was implemented directly using this technique.

- **Post-processing**

The accuracy of an OCR can be easily enhanced if its output is being constrained by a specific list of words that can occur in a specific document. For example, that might be the words of the English language, or a more technical list of words designed for a specified area. This method cannot be of much use in case the specific document comprises of the words that are not in the lexicon (like proper nouns). Its output can be plain text file of characters, but more sophisticated systems using OCR have the ability to conserve the actual layout of the page and produce it, as in case of a marked PDF document that comprises of the original text image of document and also a searchable textual representation of it. The accurate information of the grammar of the language that is to be scanned helps in determining that if a word is likely to be a noun or a verb; thus increasing its accuracy.

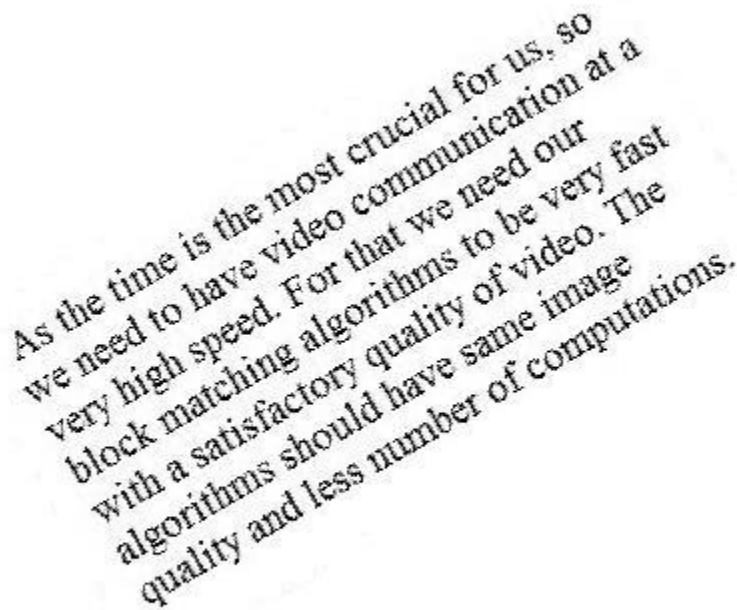
## **1.5 Advantages of OCR**

As a printed page is in machine-readable text form, we can search within it by keyword, edit it, compress it into a ZIP file, incorporate it within a web page, send it by e-mail and also store it in much less space. The machine-readable text can also be decoded with the

help of tools that use speech synthesizers for reading the words on a screen so that the blind and visually impaired people can understand them easily. Earlier one of the major use of OCR was in a device that were able to read printed books loudly for the ease of blind people.

## 1.6 Skew Estimation and Correction

The presence of skewness in any of the scanned text documents leads to the failure of layout analysis algorithms used for analysing the documents. Skew angle is actually the angle that is made between the text lines of digital image and horizontal direction of the text document. Skew estimation means determination of the angle at which scanned text document is skewed (or tilted) from the normal position. After determining the skew angle of the skewed text document image as shown in Figure 1.4, a skew correction technique can be applied on it to obtain the final de-skewed image as shown in Figure 1.5.



As the time is the most crucial for us, so we need to have video communication at a very high speed. For that we need our block matching algorithms to be very fast with a satisfactory quality of video. The algorithms should have same image quality and less number of computations.

Figure 1.4. Skewed text image

Skew correction technique can be used for de-skewing the scanned computerized text document, handwritten text document, and document written by a typewriter. Skew

correction and skew estimation are main pre-processing steps used for the document layout analysis and also for optical character recognition approaches.

As the time is the most crucial for us, so we need to have video communication at a very high speed. For that we need our block matching algorithms to be very fast with a satisfactory quality of video. The algorithms should have same image quality and less number of computations.

Figure 1.5. De-skewed text image

Skew correction is a process of first detecting whether the document has some slope/tilt or not and then rotating that document if the angle of slope is significant, so that after skew correction the baseline of that document becomes horizontal. It increases the performance of the system [3]. The slope/tilt is the angle made between the direction of line on which the document is aligned and the horizontal direction. If document is scanned manually or mechanically then some degree of skewness is unavoidable and it results in a skewed image as shown in Figure 1.7 that represents skew of only 7 degrees.

**After you type the text, make the Free Transform tool active. With it, click and hold on the top center "handle" of the text box. While holding the mouse button down, press and hold Command (Mac) or Control (PC) on your keyboard. Until you hit this key, you would only be able to drag up and down. With the modifier key held**

Figure 1.6. Original image in correct alignment.

**After you type the text, make the Free Transform tool active. With it, click and hold on the top center "handle" of the text box. While holding the mouse button down, press and hold Command (Mac) or Control (PC) on your keyboard. Until you hit this key, you would only be able to drag up and down. With the modifier key held**

Figure 1.7. Original image skewed by 7 degrees.

Generally, a little skew of 0.1 degrees or 0.2 degrees is easily perceptible for human beings. A reliable technique could be used for processing scanned documents either at the pre-processing stage or skew estimation and correction can be utilised. Considering an ideal case, if an input fed to the scanner is as that shown in Figure 1.7 would produce an output image as that shown in Figure 1.6.

The detection and correction of skewed text documents are important issues in realization and application of practical document reader. Most of the systems using Optical Character Recognition (OCR) are much sensitive to the skewness present in the scanned text documents. Still many methods used for OCR systems do not give hundred per cent precise results because of problem in detecting the accuracy of skew angle.

**LITERATURE REVIEW**

---

**2.1 Skew Detection and Estimation based on Hough Transforms**

The skew estimation by using Hough transform [3] is a widely used approach used for machine vision. Regardless of being computationally expensive, it has been widely employed for the detection of lines and curves present in the digital images. The Hough transform approach considers the point, that there are high numbers of pixels present on the lines that are coincident with the baseline of the text [3]. For improving the efficiency in respect of computations in the Hough transform-based approaches, many variants of it have been proposed that efficiently and effectively reduces the number of points that are mapped within the Hough space. This can be easily achieved by selecting a subset of pixels or by restricting its analysis within a small specified region of the image. Initially, Hough transform-based document skew detection approach was proposed earlier in 1989.

The Hough transform was calculated at all the angles of  $\theta$  between 0 degrees and 180 degrees. A set of rules are employed that are used for measuring the amount of change in the values of accumulator for particular value of  $\theta$ . The angle that predicts the skewness present in an image is set to the value of  $\theta$ , maximizes the set of rules that are intended to increase the probability of solving problem. The Hough transformation technique and run-length encoding are fused together [4] for the estimating and detection of the skew angle in the document analysis and interline spacing that is a necessary parameter for the automatic segmentation of text from graphic images. In addition to that the reduction of data has been done by using the horizontal and vertical run-length computations. An image of a document that is acquired at 600 dpi, is then sampled by a factor of four and is further transformed into a burst image. The burst image is made by swapping each of the vertical black run with that of its length as placed in the bottom-most pixel of run. Additionally, the Hough transform is applied on all pixels that are present in the burst image having a value less than 25, and precisely targeting to neglect the contributions of non-textual components available.

A bin having the extreme value in Hough space gives the angle at which the text document is skewed. A data reduction technique has been introduced later in 1997 that

works directly on images in compressed form. The extraction of the points that corresponds to pass codes is made possible with a single unique pass over that of an image with reduced quality and then its mapping is done into the Hough space. This approach is further extended for images available in compressed form [5].

In [6], within a selected region they have used the lowermost pixels of that particular object which is to be considered for the Hough transform. Accordingly by finding a local peak in Hough space, the skew of the document is estimated. Based on the set of 250 text pages, analysed accuracy is of 0.5 degrees. When text becomes light, it is becomes a little difficult to select a peak in the Hough space [7].

In [8], two different techniques are purposed that are able to estimate the skewness present in the text. The first technique performs data reduction that starts from the bounding boxes of respective connected components. The main idea behind this is to eliminate the components that make contribution for noise that includes ascending and descending components, punctuations. The minor components present are filtered away together with the components whose heights are above the average heights. Two set of points L1 and L2 are then gathered that respectively contains the left-most pixel of uppermost run and right-most pixel of lower-most run of each particular component. The points within the L1 and L2 are implemented in Hough transform technique. In order to reduce the complexity of computation of approaches based on Hough transformation, here a three step algorithm has been purposed which could be effectively used for document skew detection and correction [8]. In this technique, first the document is splitted into several small blocks, and each block of it consists of hundred rows. Then the points of every block that are to be detected are recorded which will be used for the estimation of skew in coarse form. Afterwards by selecting a peak within the Hough transform in neighbourhood of coarser skew, the accurate skew is obtained.

A new has been proposed in [9] for the estimation of document and the page orientation (landscape and portrait). The detection of page layout and orientation within a page is done by further splitting the image into squares of smaller size and each of them is classified on the basis of content of their textual and non-textual data in accordance to various heuristics that considers the distribution and density of black pixels. Then each square having text within it is classified as landscape or portrait by analysing its vertical and horizontal projection profiles. This classification is subjected to the presence of the

peaks alternated with valleys and also depends on the comparison of profile variances. The total number of black pixels present in each square with text is used as a classification score. The initial level of pyramid is constituted by these squares; then each succeeding layer of pyramid comprises of the increasingly large squares, that are made by merging the groups of nine squares of previous layer. So the whole page is represented by the top of the pyramid. The available information about the classification is being transferred from the base of the top as each square is categorized as landscape or portrait by a widely held criterion among that of the nine underlying squares, considering the scores needed for classification. The skewness is then predicted based on the sub-region of the image that corresponds to that of the highest recorded square amongst the nine last layer of the pyramid. Skewness can be calculated by implementing the Hough transform on black pixel of last row of each of the component that is being connected.

A fast and accurate approach for low resolution images have been introduced in [10]. This technique is based on the fusion of Hough transform and the connected components with their centroids. This fusion method is suitable to be used for de-skewing the images of handwritten text, forms, postal labels, drawing and bar-codes. Firstly this algorithm efficiently computes the connected components and also computes their centroids. These components are connected here by a arrangement called a “block adjacency graph”. Using two angular resolutions, the Hough transform is applied on centroids. The resolution permits an estimation of skew angle and it also restricts the angular range at which the Hough transform should be computed at fine resolution.

In [11], the Hough transform is applied for efficiently removing the skewness present in the image in two different stages. The text written together such as captions of pictures and paragraphs are acknowledged first. Then by fitting the straight lines with the help of least square method, the skew angle for each block is calculated. For boosting the speed only the lowermost line of a block is to be considered for detecting the amount of skewness present.

In [12], based on improved Hough transformation the skew has been computed for a document image. First the black runs are smoothed by using concept of image smoothing and then determining the black-white transitions for pointing out the lines in which text is present. Some other techniques are also available for the skew detection process that are language dependent and can be applied only for a document available in a particular

language. Different strategies have been proposed by the researchers for reduction in the quantity of input data fed while using the approach of Hough transform but still their cost in terms of computation is not reduced. Before applying the Hough transform, it is necessary to extract the text regions present within the document images; yet it is still very difficult in case of scanned document images because the layouts are complex and unknown. The documents with images or photos only could not be de-skewed using Hough transform approach.

## **2.2 Skew Detection and Estimation based on Projection Profile**

Initially the projection profile approaches were only a solution for the detection of angle that gives skewness within that of a document image. It was first proposed in 1986 [13] and was based on the horizontal projection profile. A sequence of horizontal projection profiles are generally calculated at a range of expected angles. In the method discussed, a large variation has been observed for reducing its high computational costs. The main aim is to condense the amount of data that is involved in the calculation of profile or for enhancing strategy involved for optimum search. For determining the skew angle, the features are taken out from each projection based profile. The profile having the maximum variations are referred as the best alignment to the lines with text written on it. In documents having skewness present in it, the angle of projection is the angle that gives the amount of skewness present in the document.

The projection profiles shown in Figure 2.1 to Figure 2.4 are calculated from the images of Figure 1.6 and Figure 1.7. In Figure 2.1 and Figure 2.2, peaks calculated are the resultant of de-skewed image shown in Figure 1.6, and the peaks in Figure 2.3 and Figure 2.4 are the resultant of skewed image as shown in Figure 1.7. A technique was proposed for the selection of the projection points that are to be used for each connected component [15]. Its main task is to figure out the sum of squares of the profiles. For fast searching of the optimal angle for de-skewing the text, an iterative method has been proposed.

Here, the entire range of angles available is searched with coarse angular resolution in the first iteration and further in succeeding iterations the range of angle is being limited to a specific neighbourhood of best angle present in current scenario and the resolution of the angular range is increased.

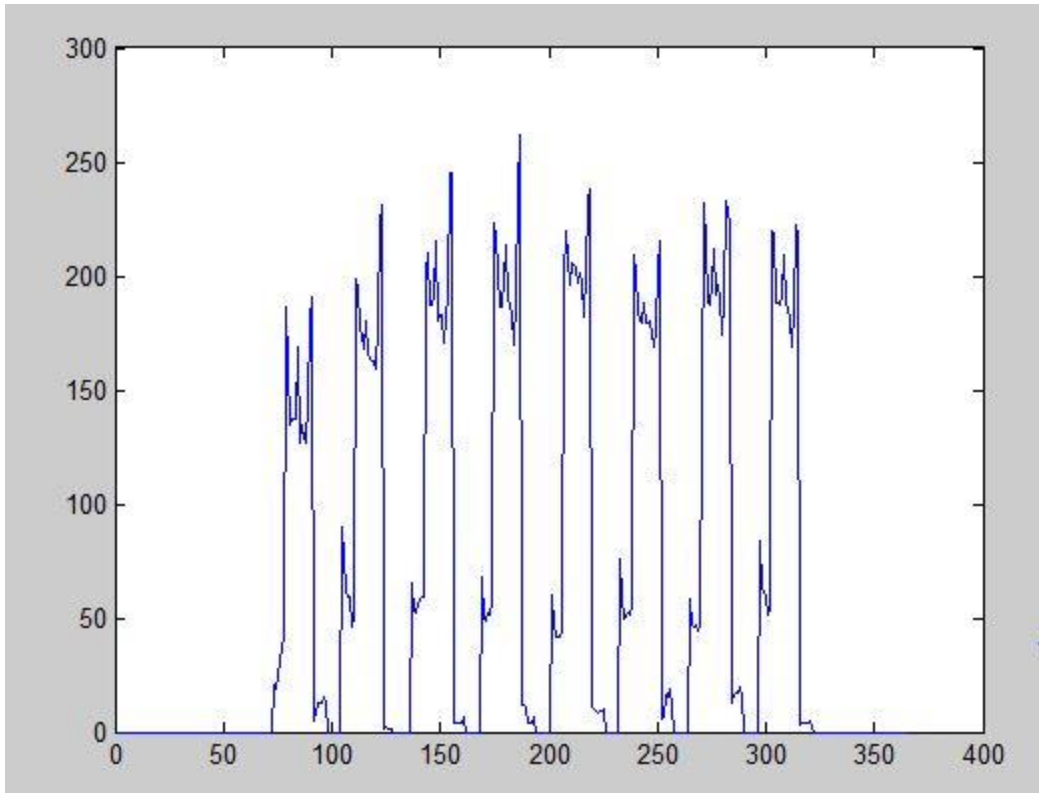


Figure 2.1. Horizontal projection profiles of de-skewed image.

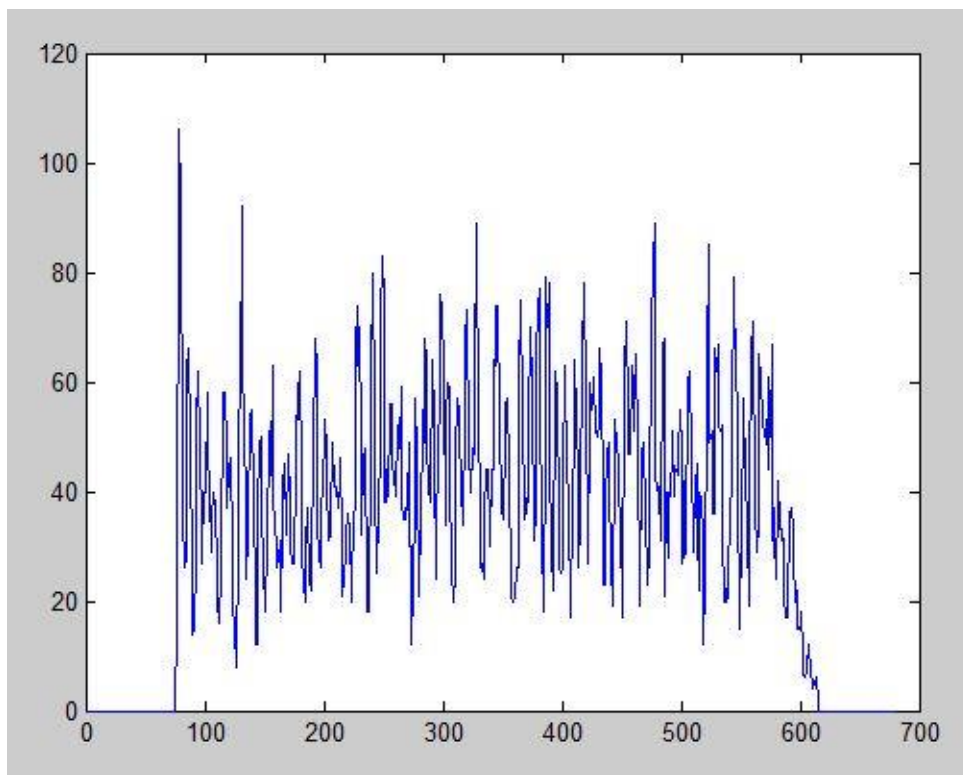


Figure 2.2. Vertical projection profiles of de-skewed image.

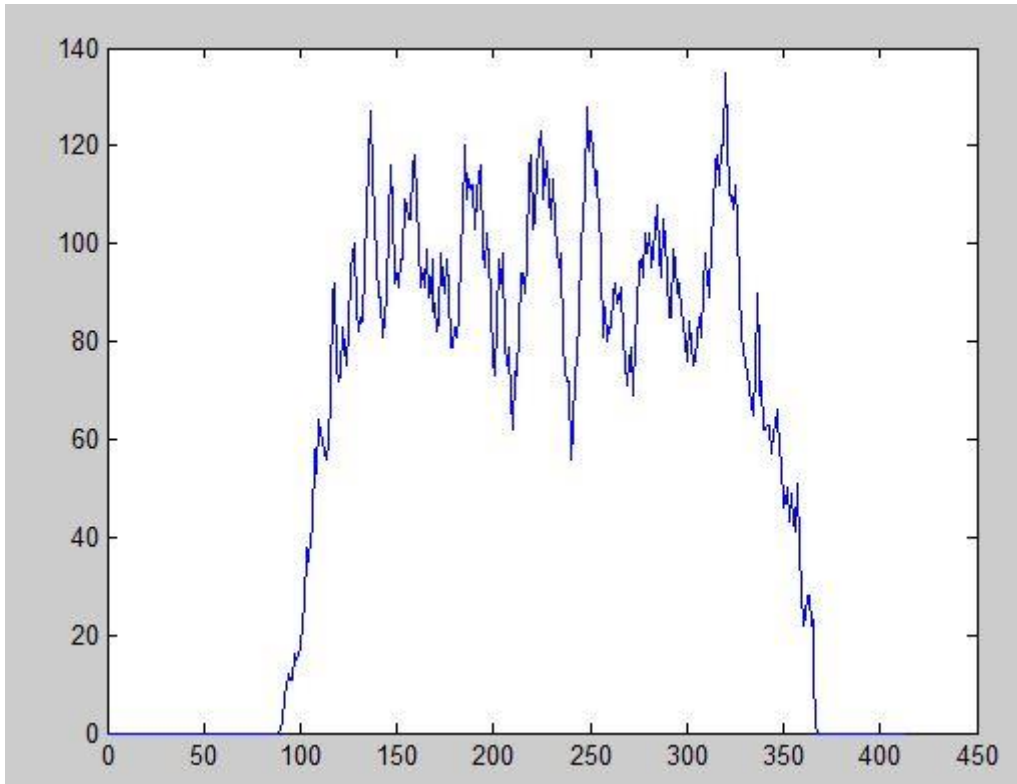


Figure 2.3. Horizontal projection profiles of skewed image.

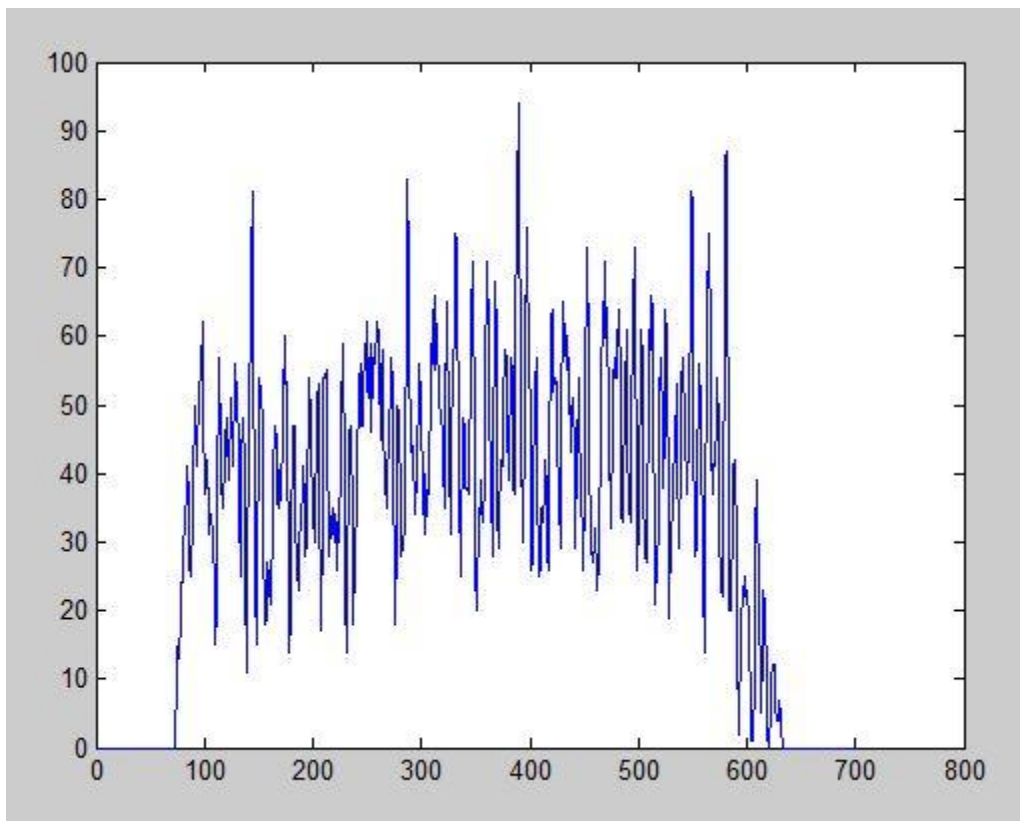


Figure 2.4. Vertical projection profiles of skewed image.

Then in further years of research, a fast and accurate approach [16] was claimed that was implemented without any modification on variety of layouts comprising of sparse tables, multiple available blocks, mixed document page sizes. Later it is defined by using a profile that is defined in different styles. Here, the number of white/black transactions along with the lines is stored in the bins of the profile and a cluster of the lines that are parallel on image is selected. Its objective is to maximize the variance by changing the slope of lines. In the presence of non-textual regions, this is a robust method.

In [17], a method is proposed in which Fourier transform and projection profile are fused for de-skewing the skewed document. A high accuracy is there even for documents with photos and graphics.

In [18], a technique has been proposed for choosing points from that of a document image that is present in compressed form. Its detection can be done easily by analysing a bit stream present in compressed form. The detection of White Pass codes are done by simulation of a process of decoding by the use of two-state automation. Few other researchers calculated the angle that gives the skewness of the scanned document image. By integrating the outcomes of a series of vertical projections of documents, a document image can easily be segmented into columns. The centres of the projections that span the particular column are predicted and least square that fits to their centres are further calculated. The skew of that specific column is the slope of the best fit line. The skew angle of the full page can easily be determined by simply averaging the skew values of the columns. It has been proved that the horizontal projection profile is computationally fast approach than that of the vertical projection profile approaches for comparison. To enhance the speed, the computational complexity of the approach that is based on projection profile has been reduced. An original document image is reduced accordingly by extracting a sample image from that of a skewed document.

In [19], the skew angle can easily be calculated by that of a sample image rather than by the whole document. Yet there are some demerits observed in these approaches. First the document analysis layout influences the accuracy in the detection of skew angle of the document image. Further for those document images that contain various fonts, non-textual regions as that of pictures, graphic images or charts, and different character sizes, it does not bring accurate results for them and variations are very big.

In [20], a method has been introduced for the correction of the amount of skew present within the document and is independent of the contents of the document. The document image is further divided into text/non-text blocks and is based on the features extracted from the projection profiles and Fourier transform of the blocks. For the determination of skew only for text blocks, these features are fed to Support Vector Machine (SVM) by considering the standard deviation in the projection profile for various angles. The angle having the maximum standard deviation is considered as the skew angle of the skewed document image. The approaches based on projection profiles are computationally expensive because in a particular range, different projections are calculated at various angles. These projection profile methods are restricted for the calculation of skew angle within  $\pm 10$  degrees. Additionally it was also emphasised that some approaches that provides the high accuracy and their range for the detection of skew angle is limited to  $\pm 15$  degrees. Accuracy mainly depends on the angular resolution of projection profiles. The projection profile based approach cannot deal efficiently for the documents containing noisy and character fragments.

### **2.3 Skew Detection and Estimation based on Nearest Neighbour**

This approach aims at exploiting the assumptions that the characters in a line are close to each other and are aligned. They are generally characterised by a process that starts from a set of objects, point representative of them or connected components, and utilizes their spatial relationship and mutual distance for the estimation of the skew angle of the document. Some researchers prefer this technique in their approach for skew removal in a document image [21] [22].

In [23] in a document image for each connected component, they pointed out five nearest neighbour. For the estimation of the skew angle of the document, in a histogram the angles between nearest neighbours are collected. The skew angle of the document here is the angle with the maximum value in the histogram. Its accuracy is for a resolution of 0.5 degrees. Using contour they fixed the boundary of the character present in the text lines [1]. With regard to the horizontal axis, the calculation of direction vector of text line is done from the growing boundary of the character till the time it reaches a pixel value of adjoining character. In present scenario, it affects the direction of growing boundary.

In [24] this technique of skew angle estimation is used mainly for the detection of dominant skewness that is present in the document image. This chain based approach has enhanced accuracy and capability of being independent of any language.

In [25] the algorithm starts by boxing the each black pixel block using the component labelling concept. It has an angle detection range from 0 degrees to 36 degrees and also can be implemented on complex text documents with a precision of 0.1 degrees. To conclude this approach is a process whose all techniques and methods start by labelling the components for grouping together the black pixels in blocks. Then the similar characteristics blocks are grouped into larger components as in case of text lines. Further, they tried for the estimation of skew angle based on text lines that are formed. This is a slow technique due to the involvement of component labelling phase that is having quadratic time complexity. Due to the grouping of lines in any direction, the accuracy of this method is reduced because of the connections with characters having noisy sub-parts.

## **2.4 Skew Detection and Estimation based on Log-Polar Transform**

This new technique for the estimating the amount of skew present in the text is based on the co-operation of two methods namely cross-correlation and log-polar transform method [26]. First step is to convert the two images (image with referent text skew and original image) into log-polar space. Then in demand to extract their similarity both the images are cross-correlated in log-polar domain. The cross-correlation of both the images gives the cost function. The cost function estimated here represents the measured amount of match found between the images [27]. The maximum value of cost function gives the optimum matching.

The angle of image rotation or say the skew angle of the image is represented by the horizontal axis of the function. By this approach the skewness present in the image is identified by its maximums. This technique is computationally inexpensive and gives a good image resolution of 300 dpi. The better results obtained using this technique is a result of the binary image processing. The points in the Cartesian space are mapped with the points in the log-polar parametric space from the image, using this transform. In the image, depending on the position it reduces the data vector. This method is also appropriate for the documents with printed text. This method is more accurate in detecting the angle of skewness present in the scanned document image as compared to that of the previous methods.

## **2.5 Skew Estimation based on Interline Cross Correlation Clustering**

The aim of such techniques is to estimate the angle that gives the amount of skewness present in the document by measuring the horizontal and vertical deviations along the

image. In these approaches, it is assumed that de-skewed images with only text, present a same kind of horizontal structure. A technique in which document is divided further into vertical blocks having same width [28]. Here, the horizontal projection profiles of each block are computed with the shifts and that shifts provides the linking of each projection with next one. The angle of skewness within a document can be determined by the equation (1) as given below.

$$\theta = \tan^{-1}(\text{mean shift} / \text{block shift}) \quad (1)$$

In [29], they have estimated the skew angle by determining the cross-correlation between vertical slices within a document and in the cross-correlation function finding a peak. Here, the efficiency is reduced by introducing the charts and graphics within the text. In [7], a method for the estimation of angle of skewness in a document has been developed based on the linear regression. This approach is appropriate only for the documents available in printed form, and for de-skewing the documents having skew angle greater than 30 degrees this method fails. For obtaining the document skew angle, the cross-correlation is calculated for the entire image. If graphics are present in the image, the accuracy is degraded and it becomes a more time consuming task [6].

## **2.6 Skew Estimation based on Fused Techniques**

In [14], based on the projection profile analysis and wavelet decomposition technique a skew detection algorithm has been proposed. First the document images present in skewed form are divided into the sub-images using wavelet transform. The matrix that contains the mean values of horizontal sub-band coefficients is rotated within a range of angles. Then at each angle within the defined range of angles a projection profile is computed. Here, the skew angle is defined as the angle that maximizes a criterion function. Such fused technique uses more memory for processing and is highly complex to implement. While using this technique, some authors considered that space between texts lines present with text should be greater than the space between intra-word and inter-word, without considering the fact that in real world applications it is not mandatory.

In [13], the Fourier transform of entire scanned page of document is being computed and it uses the power spectrum of the Fourier space for associating a score with each angle that is designated for selection. Here, skew angle is the angle that maximizes the score.

In [30], the angle that gives the amount of skewness in the scanned document is detected using the gradient direction analysis, applied to greyscale or binary images. In order to map the magnitude and direction (or say for gradient mapping), the image is divided and convolved with two masks. For each cell of a grid, the local dominant direction is computed by using the information about its magnitude and direction. The histogram of such directions can be figured out after angle quantization. The text direction can be estimated by using the maximum peak in the histogram.

In [31], a technique has been proposed assuming that in a scanned document image there are few additional points that have their gradient orientation perpendicular to the lines in which text is present. Then the computation of the histogram of the gradient orientation of greyscale image is done. In an effort for reducing the undesirable effects which arouse due to process of quantization, the smoothing of histogram is done using a median filter. The angle that predicts the skewness can be estimated by the mode of the histogram. The bounding box technique developed in [1] is suitable only for globally skewed documents and for tracing the box the corners are identified, and after the box is developed its angle of skew is estimated with the horizontal line. The approaches discussed under this category are a mixture of all the categories discussed above. Fused methods enhance the accuracy of the system and the memory requirement, computational complexity and demand of processor are ignored.

## **2.7 Skew Estimation based on PPA**

A skew detection and estimation algorithm has been developed here known as Piece-wise Painting algorithm (PPA) for scanned document images [32]. Two painted images can be obtained on applying the PPA algorithm in horizontal and vertical direction. With the help of some statistical information the regions with a particular width (or say height) from the vertical (or horizontal) painted images are being selected and points on the left on top side, middle, and the right on bottom side of such regions have been grouped into six separate lists. Using the concepts of line drawing and line regression, from the lists of the points, few lines are obtained and then their individual slopes are being calculated.

In [33], a technique is developed for finding the best fit line amongst all the lines and hence from the best fit line the skew angle of the input document image is estimated. This technique of skew estimation is independent of the writing mode of the document. In this

method it needs not to rotate the whole image, as that is a computationally expensive task. This algorithm yields better results both in case of accuracy and computational time.

## 2.8 Skew Estimation based on Image Dilation

In this research work an efficient technique for detection and estimation of amount of skewness present, has been proposed that is used in estimating the angle which gives the skewness in document images present in binary form. Using a line as a structuring element, experimentally whose length is fixed, the input document image is being dilated with it and using depth first search the region labelling technique is applied. Then for all the labelled regions the orientation angle is calculated and the angle of skewness the scanned document images here is the average of all the orientation angles. From the results shown here we analysed that improved estimation accuracy can be obtained using this particular approach as it depends on the orientation angles of text lines of the considered document. This method is vigorous for the documents that are printed by a machine, document can be of any font and size; and also for scanned document images that contains pictures, charts, graphics, tables etc.

## 2.9 Skew Correction Methods

The skew of a scanned document can be corrected by rotating the skewed document by the estimated skew angle. The methods used for the correction of document skew can be broadly classified into three categories as mentioned below:

- (1) Direct methods
- (2) Indirect methods
- (3) Contour oriented methods

### 2.9.1 Direct Methods

The image can be de-skewed using the direct methods by rotating black pixel by  $-\theta$  as in [10]. As explained, in the input image the black pixel  $b$  is changed to  $b'$  by multiplying the coordinates of pixel  $b$  by a rotation matrix as shown below in equation (2).

$$\begin{bmatrix} a' \\ b' \end{bmatrix} = \begin{bmatrix} \cos(-\theta) & -\sin(-\theta) \\ \sin(-\theta) & \cos(-\theta) \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} \quad (2)$$

where  $(a, b)^t$  are the coordinates of  $b$  in the scanned input image and the coordinated of  $b'$  are  $(a', b')^t$ . Yet it is a fast and simple transformation but restriction of this approach is the problem of rounding in discrete plane. Hence in the resultant image, the isolated components are split and the neighbouring connected components are merged. The image gets distorted as the document after its skewness is being removed suffers from the illegal holes. Few problems are caused in the processing of documents due to this undesirable transformation.

### 2.9.2 Indirect Methods

The complement of direct methods is known as indirect methods [13] [15]. In the output image, for a pixel  $(a', b')^t$ , the indirect methods find the corresponding pixel  $(a, b)^t$  in the input image and then a value  $(a', b')^t$  is to be set with that of  $(a, b)^t$ . Applying the inverse rotation matrix, the correspondence can be computed as per equation (3) as shown below.

$$\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \cos(\theta) & -\sin(-\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} a' \\ b' \end{bmatrix} \quad (3)$$

The indirect methods are computationally very expensive because they consider all the pixels present in the output image but these methods solve the rounding problem. So, the output of these methods is definitely better but is still more time consuming.

### 2.9.3 Contour Oriented Methods

Contour oriented methods can be applied for the detection of contours of the connected components and for de-skewing those components. All of the connected components get traced with their contour string. The description of the boundary of the connected component is the contour string. This method is much efficient because of the reason that it rotates the relevant information only; and here the contour string should be closed manually always to avoid the rounding mistakes. A rotation for a connected component is a three step process here.

In the rotated image, the location of the start corner is computed first. Then the each contour edge present in the contour string is rotated. Finally it rebuilds the contour string as it finds the new and final position of the start corner.

## **2.10 Skew Correction using Cyclostyle Matching**

In this work, a real time and robust technique for skew-detection is presented that is based on the grey projection cyclostyle matching of document image and is used for practical applications. This algorithm is based on analysis of the grey scale projection curves in vertical and horizontal direction of the images that are considered for implementation, and compares with all of them if it is not homology.

It further solves the problem of finding the accurate skew angle. By using the dual linear interpolation algorithm, the problem of image distortion can be solved with help of the skew adjustment algorithm. The skew angle can be detected and corrected by detecting the markings in the image. In order to reduce the computations, while correcting the skewness within an image document, the cross-grey projection cyclostyle matching algorithm is used. This process can be realized easily by implementing the following steps [34] :

- 1) The pre-treatment of image and conversion of image from RGB to greyscale.
- 2) To detect the skewness in the scanned image.
- 3) The adjustment of skewness in the scanned image.

Finally, in the experimental results it has been illustrated that this algorithm enhances the speed of cyclostyle matching and can solve the problem of consumption of more time that arise during the skew detection and skew adjustment in case of a scanned document image and also in case of other image processing; identification of notes, cards, etc.

### ANALYSIS AND METHODOLOGY

---

#### 3.1 Introduction

Skew estimation is a process of detecting the deviation of document orientation angle from that of horizontal or vertical direction. It is assumed by the automatic document recognition systems that all the printed documents have same text direction and, therefore, there is no need of skew estimation or correction in it. This supposition was made untrustworthy and the skew correction and estimation became unavoidable as the frequent applications of flat-bed scanners to scan large numbers of documents.

Sometimes page skew is inevitable during the process of scanning or copying. Today it is a common feature that documents are even more free styled and text aligned along with the different directions. Skew estimation in documents is of two types; global and local. The global skew estimation is most popular in the research community and many other techniques have been proposed to correct it but especially for the documents with charts, figures, and graphics or different font sizes it is still an interesting but challenging issue [37]. On the other side, some researchers have developed the techniques for local skew estimation. In local skew estimation an additional page-layout analysis stage is to be included in for the accurate document restoration. In terms of efforts and memory usage it is difficult [38].

The global skew estimation approaches are further classified into four categories [39], based on the approaches they adopt. These categories are Hough transform, projection profile, nearest neighbour clustering and cross correlation.

In case of printed text documents, the need for automatic character recognition is increasing because a huge amount of data is available in it and for sorting and updating that huge amount of database; for automatic document delivery; document archiving and other applications it requires processing. Image recognition is an off-line process which means that the printed documents are scanned as images and stored. Then further in a later stage the character recognition is performed on that scanned document image. While the operation of scanning of a document is being performed, there are some documents (papers) that are not properly fed to the scanner; initiating a skewness (or say tilt) in the scanned document image.

## 3.2 Implementation

First consider a handwritten or computerized text document. Then scan the document, set the resolution of scanner at 300dpi or above and save the output scanned document image. Now a scanned document image has been obtained. The scanned document image may have some skewness present in it due to the human error while scanning. For automatic removal of skewness from an image, a skew detection and correction algorithm has been developed here.

1. The scanned document image is in RGB format will be converted into a greyscale image for further processing.
2. A technique known as median filtering is applied onto the greyscale image, which leads towards the reduction of the noise present in the image. In median filtering, a window slides along with the image and the median intensity value of the pixels present within the window becomes output intensity of the pixels which is to be processed. Median filter is a non-linear filtering technique which is used for reducing the impulsive noise and salt-and-pepper noise. It preserves the useful details while reducing the random noise from the scanned document image.
3. The image obtained in Step 2 is converted into the binary image for its further processing. An image that is having two pixel colours, black and white; in which '0' represents black and '1' represents white. Binarization of image is done to perform the computations faster.
4. Colour of pixels present in the image is changed by performing the complement operation on the image. Performing this operation on the image precisely implies reversal of black pixels and white pixels where the pixels having value '0' become '1' and the pixels having value '1' becomes '0'; that now indicates the text area present within the scanned image will be represented with white pixels, and the background and edges outside of the scanned text area will be represented with black pixels. It becomes a bit important to perform the operation of interchanging the colour of pixels because of the fact that further steps of our algorithm will rotate the image, as shown in Figure 3.1, and shape and size of image changes further as shown in Figure 3.2.

As the time is the most crucial for us, so we need to have video communication at a very high speed. For that we need our block matching algorithms to be very fast with a satisfactory quality of video. The algorithms should have same image quality and less number of computations.

Figure 3.1. Skewed image

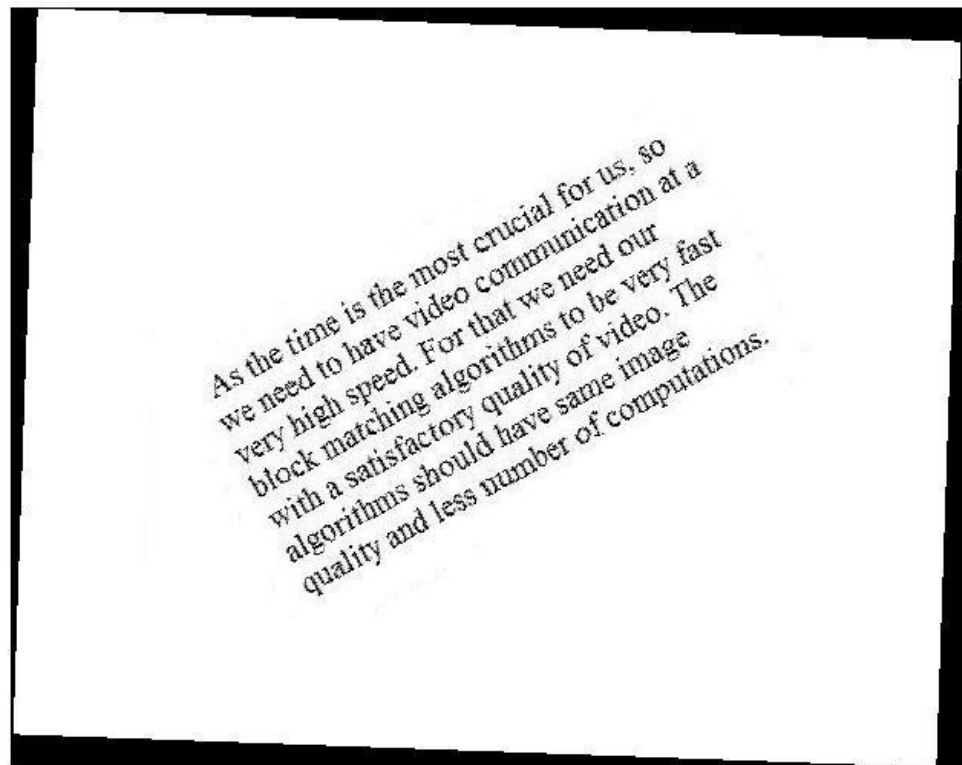


Figure 3.2 Skewed image rotated by 2 degrees

An unwanted region that is black in colour present on the boundary of the image (as shown in Figure 3.2), is introduced and it leads to the hindrance in the following operations of de-skewing the scanned skewed image. The hindrance arises because of the fact that our main aim is to identify and extract the region containing text. This can be done by applying a technique in which we clip the portion around the boundary of the area that is containing the text (please refer to Figure 3.4).

To perform the operation of clipping it requires complementing the image, otherwise when the scanned document is rotated as such, a black background on the outside boundary will appear (refer Figure 3.2) and subsequent steps for identifying and extracting the boundary around the text area will not be possible. It can be explained in other words as the first black pixel in Figure 3.2 is the very first pixel from the left hand side and in this way the starting black pixels of the text cannot be reached by simply calculating the first black pixel from each row (or column).

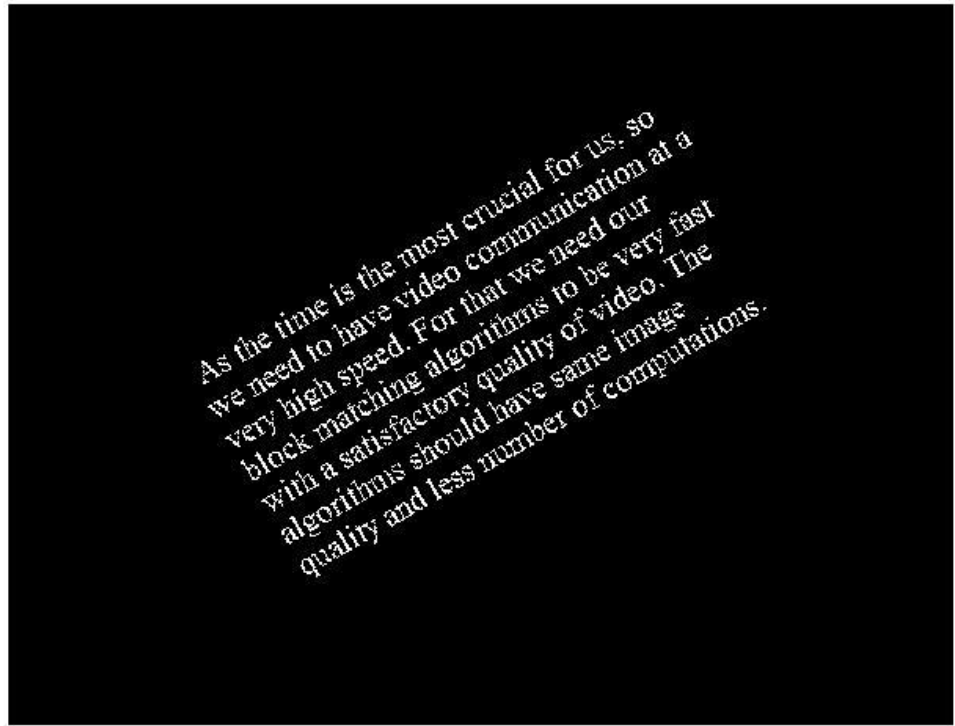


Figure 3.3. Complemented skewed image

So, finding the complement of the image is mandatory here, as complementing the image as shown in Figure 3.3 will help us in easily calculating the first white pixel (text pixel) as the area except the text is all having black pixels. The text can be extracted easily using this approach.

5. Define the text area of the complemented image of Figure 3.3 as shown in Figure 3.4. This reduction in size can be done by utilizing the information about the location of the corner pixels of the text present in it.
6. The image is cropped to the new size as that defined by the text area of the image with the boundary on its sides and denote the new image by  $I'$  as shown in Figure 3.5.
7. For calculating the de-skewed image, next step is to find the number of white pixels present in each row of the given image ( $I'$ ).

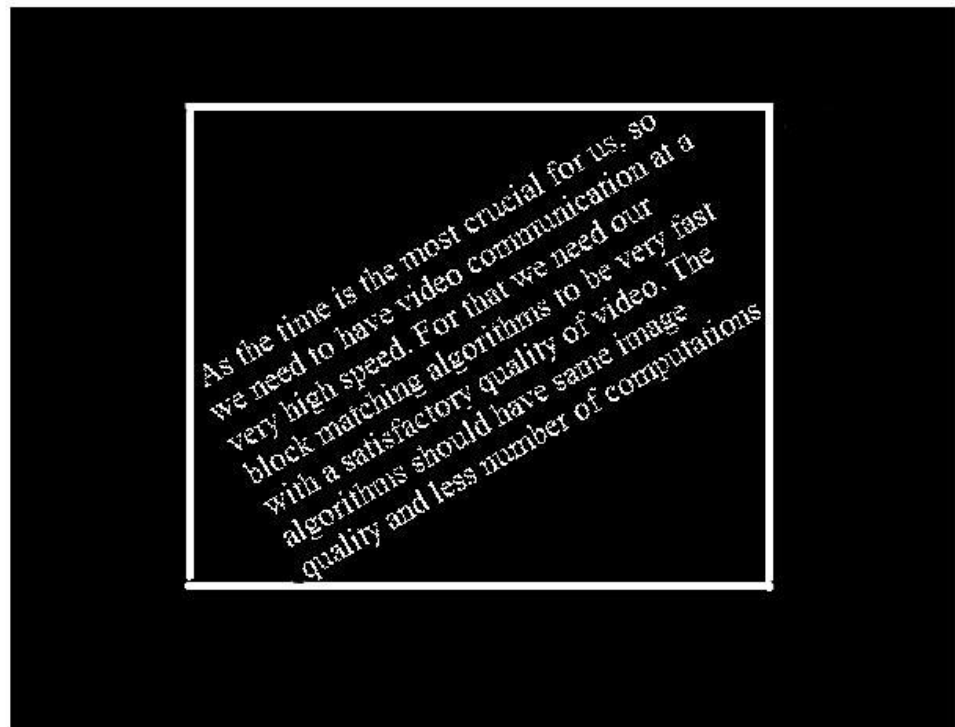


Figure 3.4. Text image within boundary

8. Now the image is rotated by a small non-zero random angle  $\theta^\circ$  in clockwise direction.

$$\begin{bmatrix} i' \\ j' \end{bmatrix} = \begin{bmatrix} \cos(\theta) & -\sin(-\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} i \\ j \end{bmatrix} \quad (4)$$

$$i' = i * \cos(\theta) - j * \sin(\theta)$$

$$j' = i * \sin(\theta) + j * \cos(\theta)$$

where,  $(i, j)$  is the pixel coordinates of image  $I$  and  $(i', j')$  is the pixel coordinates of image  $I'$ . The pixel values at  $(i, j)$  is shifted to  $(i', j')$  after rotation by angle  $\theta$  in counter clockwise direction.

We select a value of  $\theta^\circ$  in such a way that it will be able to depict the amount of tolerance that can be neglected; and the rotation will have an error of  $\theta^\circ$ .

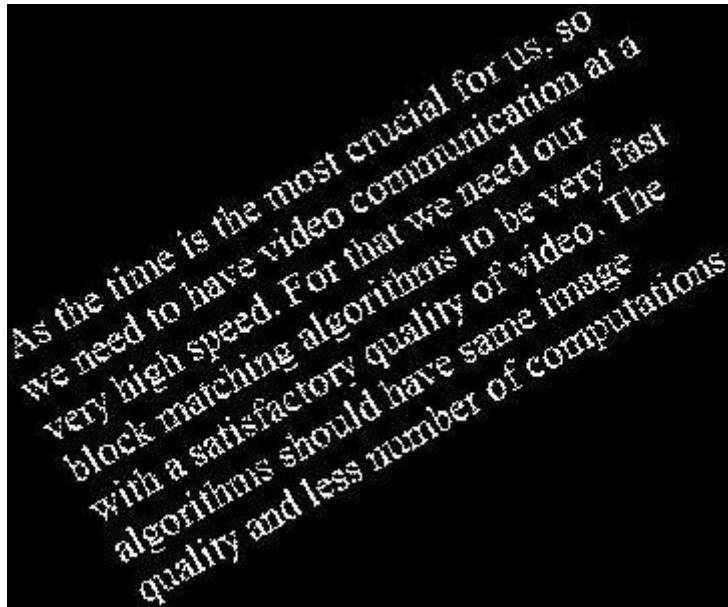


Figure 3.5. Cropped image,  $I'$

9. After rotating the image by a small angle  $\theta^\circ$  here in clockwise direction, a new image is generated with different background and size (refer Figure 3.2).
10. Remove those edges of the image that are not required for further processing as defined in Step 5 and Step 6. Here, a new image is obtained and is denoted by  $I'$ .
11. In the new obtained image  $I'$ , calculate the number of white pixels present in each row.
12. Further rotate the image  $I'$  by an angle  $\theta^\circ$  and repeat the procedure for all the angles defined within the range. Also calculate the number of white pixels present in each row for each angle.

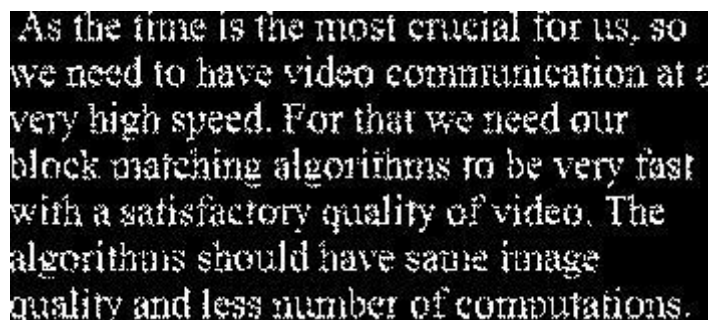


Figure 3.6. De-skewed image.

13. In this way, by rotating the image, we obtained an image in which the text lines are horizontally aligned with the horizontal plane and is known as the de-skewed image  $I''$ . The final de-skewed image  $I''$  is shown in Figure 3.6.

14. The image obtained after the mapping of pixels of Figure 3.1, is shown below in Figure 3.7 and is the resultant de-skewed image and final result of the algorithm.

As the time is the most crucial for us, so we need to have video communication at a very high speed. For that we need our block matching algorithms to be very fast with a satisfactory quality of video. The algorithms should have same image quality and less number of computations.

Figure 3.7. Resultant de-skewed image.

### 3.3 Features of the technique

- 1) **Parameter freedom:** This technique does not require any information about the pre-defined parameters such as number of text lines available within a document and many more, for finding the de-skew angle.
- 2) **Language independence:** This algorithm can be implemented for a document available in any language. It can de-skew the scanned text documents of all languages whether written horizontally (or vertically in case of Japanese).
- 3) **Angle detection range:** This algorithm can de-skew the documents skewed at any angle within the range  $0^\circ$  to  $360^\circ$ . This increases its flexibility, but if implemented with a small threshold value, the computational complexity increases a little, but will not degrade the overall performance of the algorithm.
- 4) **Layout independence:** This technique can be implemented for the scanned text document with graphs, pie-charts, tables, etc. without the prior knowledge of their position on the page and direction.
- 5) **Independence of size and font:** For de-skewing a document using this technique, any information about its size and font is not required because this technique can be efficiently implemented for the documents having any size of characters and with any font size.

RESULTS AND DISCUSSIONS

In previous chapter, we have discussed about novel de-skew algorithm, which is a generalized algorithm for de-skewing all types of scanned text document images. In this chapter, simulation results of our novel de-skew algorithm are provided after implementing this algorithm on different types of text document images.

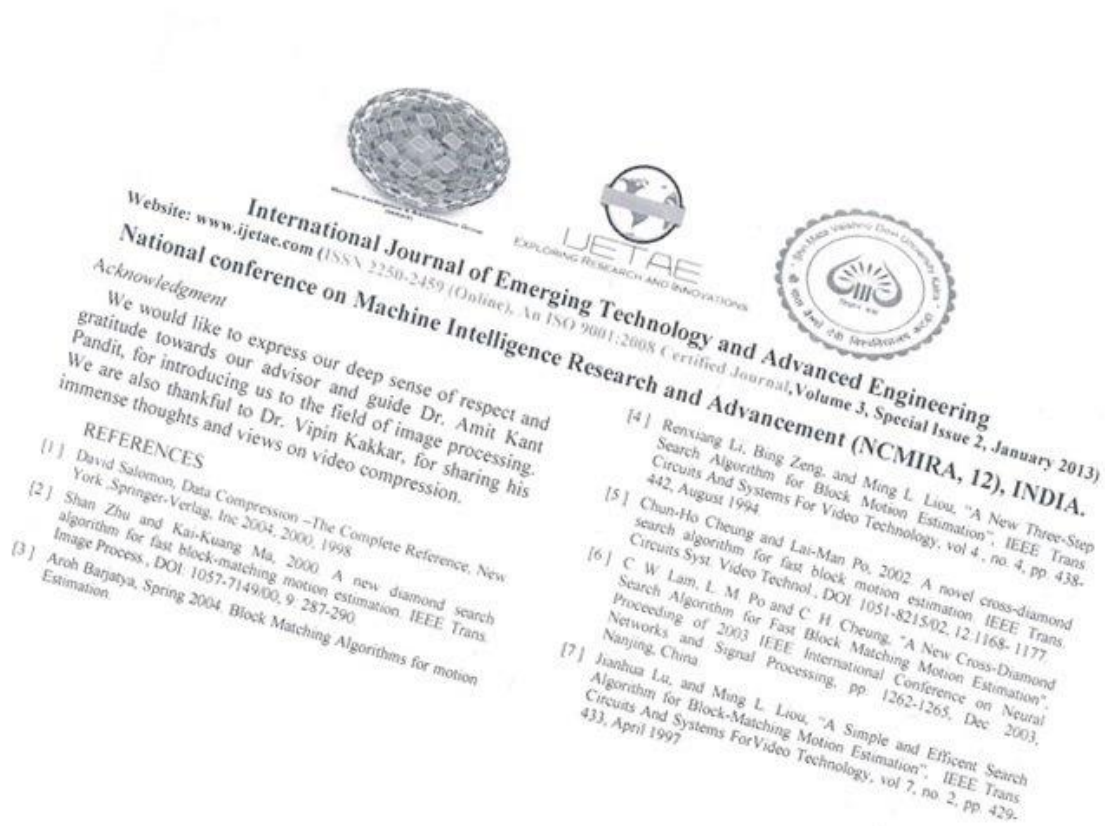


Fig. 4.1.1. Skewed document image with text and logos.



## International Journal of Emerging Technology and Advanced Engineering

Website: [www.ijetae.com](http://www.ijetae.com) (ISSN 2250-2459 (Online), An ISO 9001:2008 Certified Journal, Volume 3, Special Issue 2, January 2013)

National conference on Machine Intelligence Research and Advancement (NCMIRA, 12), INDIA.

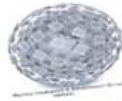
### Acknowledgment

We would like to express our deep sense of respect and gratitude towards our advisor and guide Dr. Amit Kant Pandit, for introducing us to the field of image processing. We are also thankful to Dr. Vipin Kakkar, for sharing his immense thoughts and views on video compression.

### REFERENCES

- [1] David Salomon, Data Compression –The Complete Reference, New York ,Springer-Verlag, Inc 2004, 2000, 1998
- [2] Shan Zhu and Kai-Kuang Ma, 2000. A new diamond search algorithm for fast block-matching motion estimation. IEEE Trans. Image Process., DOI: 1057-7149/00, 9: 287-290.
- [3] Aroh Barjatya, Spring 2004. Block Matching Algorithms for motion Estimation.
- [4] Renxiang Li, Bing Zeng, and Ming L. Liou, "A New Three-Step Search Algorithm for Block Motion Estimation", IEEE Trans. Circuits And Systems For Video Technology, vol 4., no. 4, pp. 438-442, August 1994.
- [5] Chun-Ho Cheung and Lai-Man Po, 2002. A novel cross-diamond search algorithm for fast block motion estimation. IEEE Trans. Circuits Syst. Video Technol., DOI: 1051-8215/02, 12:1168- 1177.
- [6] C. W. Lam, L. M. Po and C. H. Cheung, "A New Cross-Diamond Search Algorithm for Fast Block Matching Motion Estimation", Proceeding of 2003 IEEE International Conference on Neural Networks and Signal Processing, pp. 1262-1265, Dec. 2003, Nanjing, China.
- [7] Jianhua Lu, and Ming L. Liou, "A Simple and Efficient Search Algorithm for Block-Matching Motion Estimation", IEEE Trans. Circuits And Systems For Video Technology, vol 7, no. 2, pp. 429-433, April 1997

Fig. 4.1.2. De-skewed document image with text and logos.



As in case of diamond search, when the best match is found at the centre, we then switch to small diamond search. But in our rho-square search, there is no need to switch further to small search. Switching just means that we have found the best match for motion vector.

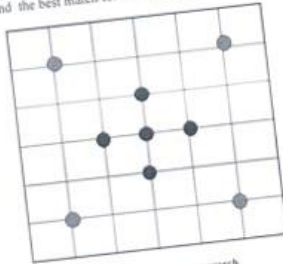


Fig.1. Rho-square search

Rho-Square method can be better implemented when the motion of the block inside the frame is to a very large extent i.e. the block is in a fast motion. As compared to the diamond search, this method can find motion vector of the required block in a lesser number of steps. Hence it has less number of computations in case of fast motion. But in case of small movement of block inside the frame, it has almost same number of computations and also the same quality of image as obtained in case of diamond search method.

#### V. DESIGN PARAMETERS

1. Computations (related to time)
  2. PSNR i.e. peak signal to noise ratio (related to quality)
- To design block matching algorithms for obtaining efficient results, we need to have low value of computations and high value of PSNR i.e. to have good quality of video compression in less time.

#### VI. SIMULATION RESULT

When the algorithms is compared with the earlier three step search method and diamond search the simulation result in Matlab is as follows:

When we compare the results of our rho square method with the three step search and diamond search, we found that the computations were reduced to a larger extent. Two tables are given in reference to the simulation result of our search method for different frames. As shown in the table1 and table2, the computations is reduced to 15.8 from 22.4, and 16.5 from 21.5, as in case of diamond search. Also there is a large decrease in the computations when compared with Three Step search method as shown in the table 1 and table2.

#### 1. From first sample frame

ALGORITHMS	PSNR	COMPUTATIONS
FULL SEARCH	40	204
THREE STEP SEARCH	39.7	23.7
DIAMOND SEARCH	39.9	22.4
RHO-SQUARE METHOD	39.3	15.8

#### 2. From second sample frame

ALGORITHMS	PSNR	COMPUTATIONS
FULL SEARCH	38.9	204
THREE STEP SEARCH	38.6	23.7
DIAMOND SEARCH	38.7	21.5
RHO-SQUARE METHOD	38.8	16.5

#### VII. CONCLUSION

From the simulation result, we can conclude that by using our algorithm, we can get almost the same quality of the video or picture i.e. PSNR, and our computations are reduced to very large extent. So finally we obtained a block matching algorithm for fast communication purposes.

Fig. 4.2.1. Skewed document image with logos, text and tables.



## International Journal of Emerging Technology and Advanced Engineering

Website: [www.ijetae.com](http://www.ijetae.com) (ISSN 2250-2459 (Online), An ISO 9001:2008 Certified Journal, Volume 3, Special Issue 2, January 2013)

### National conference on Machine Intelligence Research and Advancement (NCMIRA, 12), INDIA.

As in case of diamond search, when the best match is found at the centre, we then switch to small diamond search. But in our rho-square search, there is no need to switch further to small search. Switching just means that we have found the best match for motion vector.

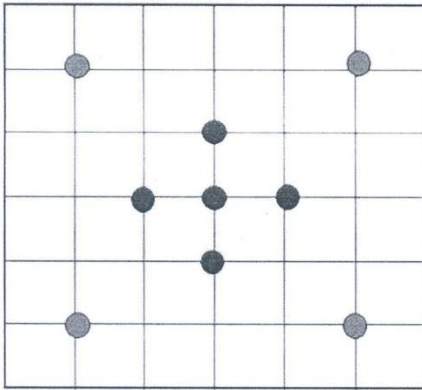


Fig.1. Rho-square search

Rho-Square method can be better implemented when the motion of the block inside the frame is to a very large extent i.e. the block is in a fast motion. As compared to diamond search, this method can find motion vector of the required block in a lesser number of steps. Hence it has less number of computations in case of fast motion.

But in case of small movement of block inside the frame, it has almost same number of computations and also the same quality of image as obtained in case of diamond search method.

#### V. DESIGN PARAMETERS

1. Computations (related to time)
2. PSNR i.e peak signal to noise ratio (related to quality)

To design block matching algorithms for obtaining efficient results, we need to have low value of computations and high value of PSNR i.e. to have good quality of video compression in less time.

#### VI. SIMULATION RESULT

When the algorithms is compared with the earlier three step search method and diamond search the simulation result in Matlab is as follows:

When we compare the results of our rho square method with the three step search and diamond search , we found that the computations were reduced to a larger extent. Two tables are given in reference to the simulation result of our search method for different frames. As shown in the table1 and table2, the computations is reduced to 15.8 from 22.4, and 16.5 from 21.5, as in case of diamond search. Also there is a large decrease in the computations when compared with Three Step search method as shown in the table 1 and table2.

#### 1. From first sample frame

ALGORITHMS	PSNR	COMPUTATIONS
FULL SEARCH	40	204
THREE STEP SEARCH	39.7	23.7
DIAMOND SEARCH	39.9	22.4
RHO-SQUARE METHOD	39.3	15.8

#### 2. From second sample frame

ALGORITHMS	PSNR	COMPUTATIONS
FULL SEARCH	38.9	204
THREE STEP SEARCH	38.6	23.7
DIAMOND SEARCH	38.7	21.5
RHO-SQUARE METHOD	38.8	16.5

#### VII. CONCLUSION

From the simulation result, we can conclude that by using our algorithm, we can get almost the same quality of the video or picture i.e. PSNR, and our computations are reduced to very large extent. So finally we obtained a block matching algorithm for fast communication purposes.

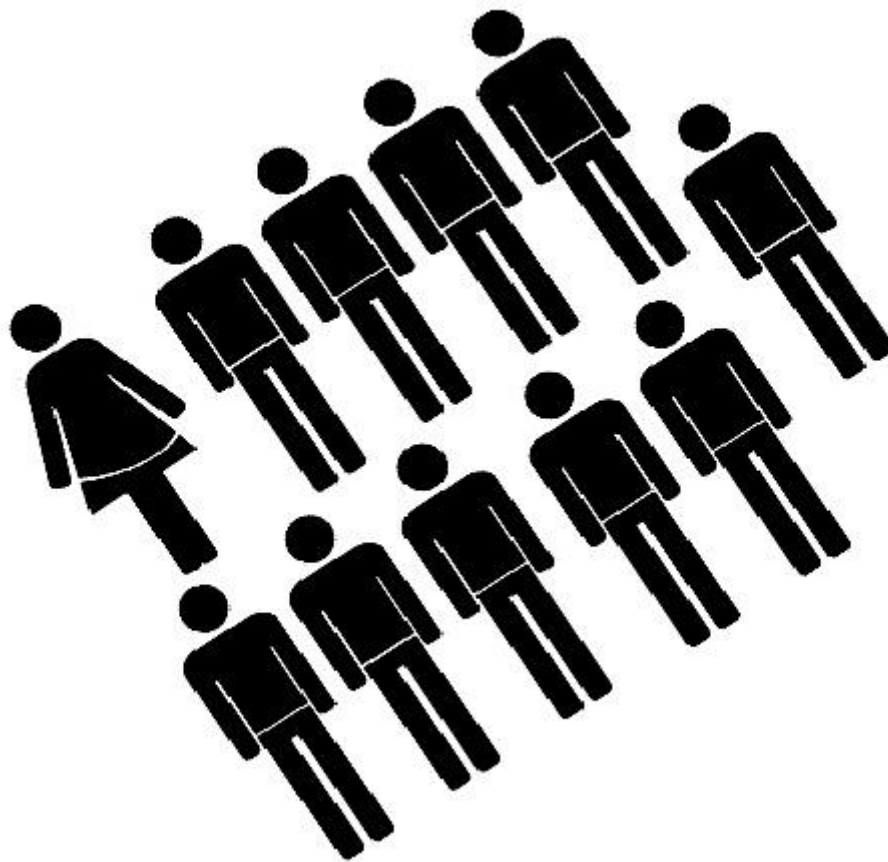


Fig. 4.3.1. Skewed document image of characters.

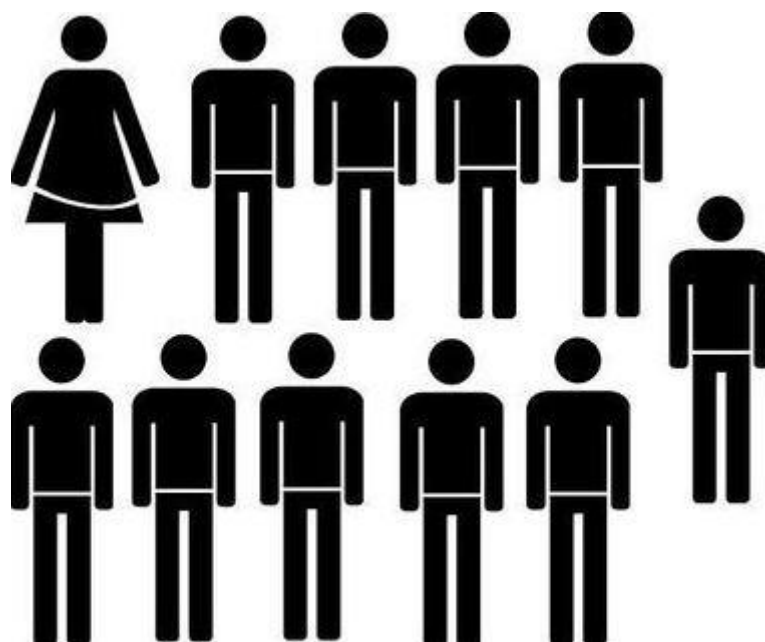


Fig. 4.3.2. De-skewed document image of characters.



Fig. 4.4.1. Skewed document having only image (with boundary).

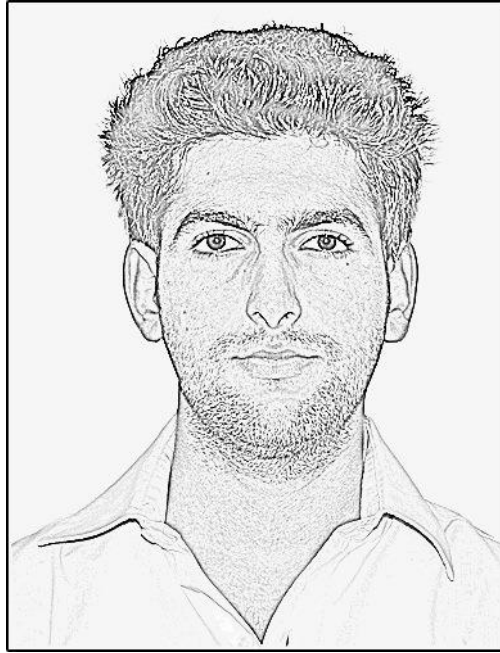


Fig. 4.4.2. De-skewed document having only image (with boundary).

तदस्मिन् जल्यमिदं वदन्ति  
शुद्धमा यदस्य प्रथमादीस्त ।  
विष्ठा यज्ञात् सहा परीवृता  
मदं सोमस्य दंडितान्यदयत् ॥  
मा नुः समस्य द्वयः पारद्वेषसा  
अंदातः ।  
आमिन् जावमा वधात् ॥  
(४.५.)

Fig. 4.5.1. Skewed handwritten text document in sanskrit language.

तदर-मं नलयसिङ्गवद-वद-यत  
शुद्धा यदर-य प्रथयादरत ।  
विष्वा यज्ञात्रा सहसा परीवृता  
मद सौमर-य दंढतान्यरयत ॥  
मा नः सस-य दुदयः पारद्वषसा  
अंढतः ।  
अमि न जावसा वधौत ॥  
(४.४.)

Fig. 4.5.2 De-skewed handwritten text document in sanskrit language.



Fig. 4.6.1. Skewed image of bar code.

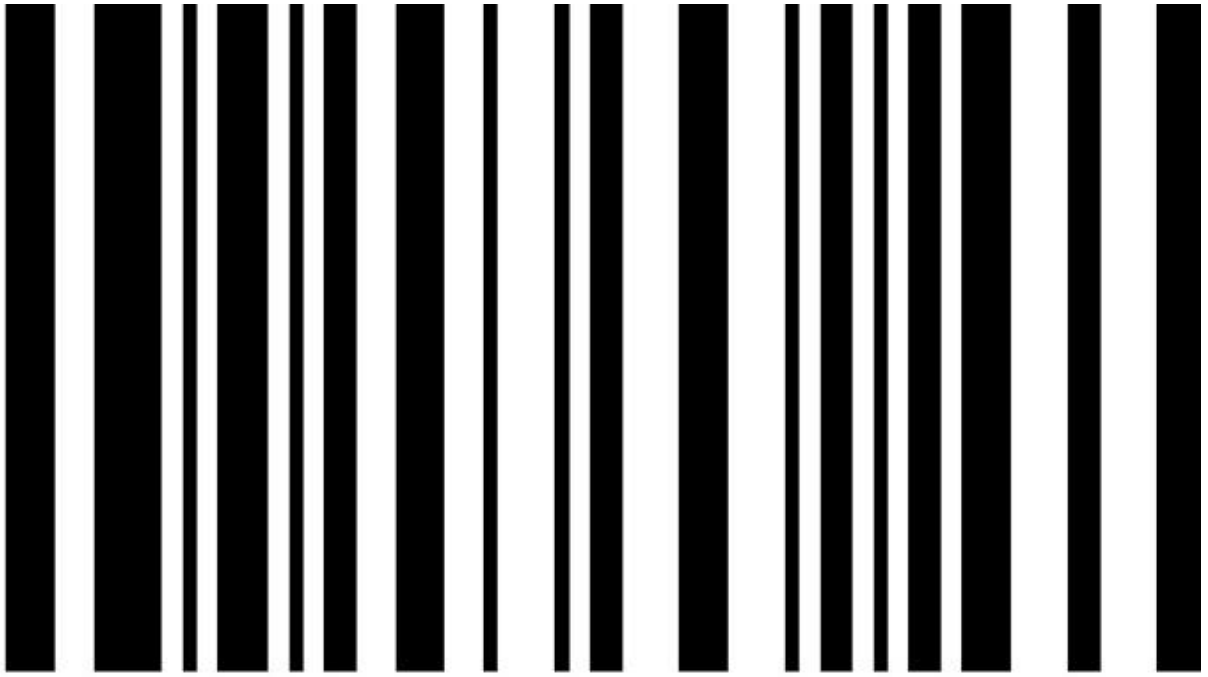


Fig. 4.6.2. De-skewed image of bar code.

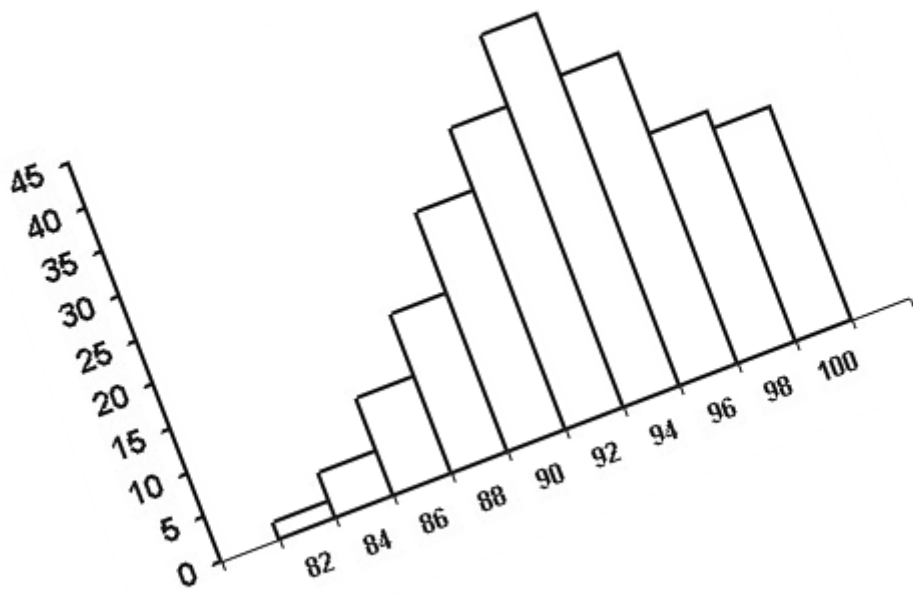


Fig. 4.7.1. Skewed image of a bar graph.

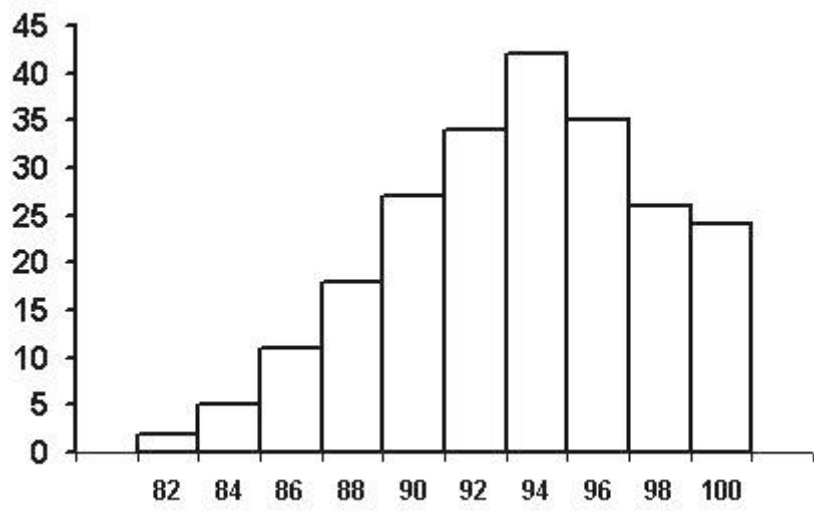


Fig. 4.7.2. De-skewed image of a bar graph.

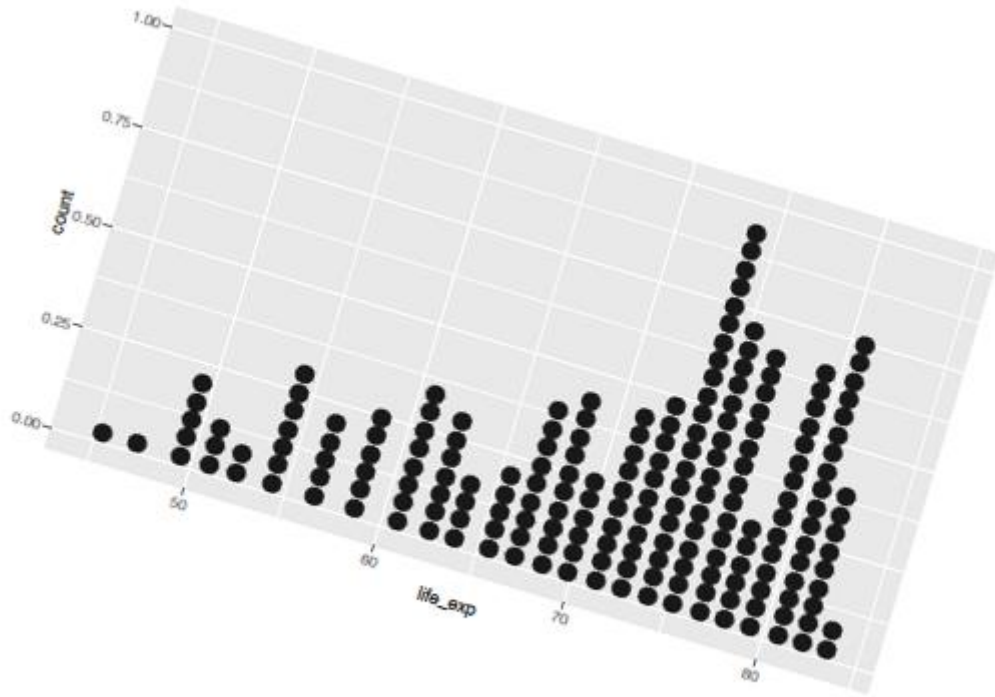


Fig. 4.8.1. Skewed image of a distribution.

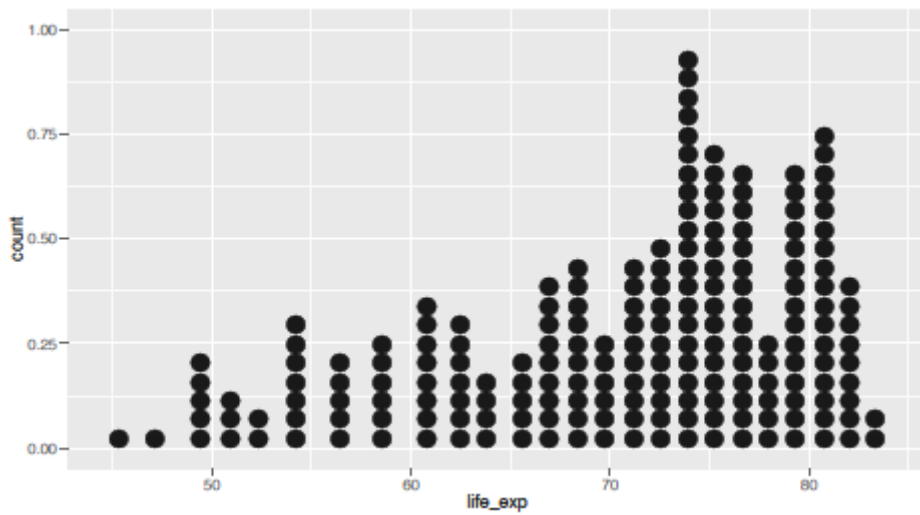


Fig. 4.8.2. De-skewed image of a distribution.

### CONCLUDING REMARKS AND FUTURE SCOPE

---

#### 5.1. Conclusion

After analysing the simulation results in previous chapter, it is concluded that the novel de-skew algorithm does not have any restriction over the detectable angle range. It is not restricted to the type and size of fonts and also does not require any specific document resolution. De-skewing a document using novel de-skew algorithm does not depends on the page layout and can also efficiently de-skew a document in which graphics, borders, and/or tables is present. This algorithm does not have a fixed computational cost. It depends on the accuracy needed to de-skew a document. If we need more accuracy, then its computational cost will increase. This algorithm is not application specific and is a generalized algorithm that can be used for de-skewing a document scanned by a scanner or is captured by a mobile device. This algorithm is language independent and can de-skew a document present in any language. It can also be implemented on bar codes, bar graphs, and distributions.

This novel algorithm can effectively de-skew the images in which logos are on the top of image and texts, tables are present within the document image. It can also de-skew the scanned document if only scanned image without text (inside a box with boundary line) is present within it. It can efficiently de-skew both handwritten and computerized text documents.

#### 5.2. Future Work

Regarding future scope, it is suggested to implement the novel de-skew algorithm on RGB scale images and analyse the simulation results.

## REFERENCES

---

1. R.C. Gonzalez, R. E. Woods, S. L. Eddins, “*Digital Image Processing Using MATLAB*,” Pearson Education Pte Ltd, 2004.
2. Joshua Doubek, “Seismic Wavelet,”  
[http://commons.wikimedia.org/wiki/File:Seismic\\_Wavelet.jpg](http://commons.wikimedia.org/wiki/File:Seismic_Wavelet.jpg), 2013.
3. M. Sarfraz, Z. Rasheed, “Skew estimation and correction of text using bounding box,”  
*5<sup>th</sup> International Conference on Computer Graphics, Imaging and Visualization*, pp.259-264, 2008.
4. J. Sadri, M. Cheriet, “A new approach for skew correction of documents based on particle swarm optimization,” *10<sup>th</sup> International Conference on Document Analysis and Recognition*, pp.1066-1070, July 2009.
5. N. Srihari, V. Govindaraju, “Analysis of textual images using the Hough transform,”  
*Machine Vision and Applications*, vol.2, pp.141-153, 1989.
6. S. C. Hinds, J. L. Fisher, D. P. D’Amato, “A document skew detection method using run-length encoding and the Hough transform,” *10<sup>th</sup> International Conference on Pattern Recognition*, vol.1, pp.464-468, June 1990.
7. A. Bagdanov, J. Kanai, “Projection profile based skew estimation algorithm for JBIG compressed images,” *12<sup>th</sup> International Conference on Document Analysis and Recognition*, pp.401-410, 2013.
8. J. Wang, K. H. Leung, S. C. Hui, “Cursive word reference line detection,” *Pattern Recognition*, vol.30, pp.503-511, March 1997.
9. P. Shivakumara, G.H. Hemantha, D.S. Guru, P. Nagabhushan, “Skew estimation of binary document images using static and dynamic thresholds useful for document image mosaicking,” *Proceedings of National Workshop on IT Services and Applications*, pp.1-5, February 2003.
10. H. F. Jiang, C. C. Han, K. C. Fan, “A fast approach to the detection and correction of skew documents,” *Pattern Recognition Letters*, vol.18, issue 7, pp.675-686, July 1997.
11. S. Daniel, R. T. George, H. Wechsler, “Automated page orientation and skew angle detection for binary document images,” *Pattern Recognition*, vol. 27, issue 10, pp.1325-1344, October 1994.
12. B. Yu, A. K. Jain, “A robust and fast skew detection algorithm for generic documents,” *Pattern Recognition*, vol.29, issue 10, pp1599-1629, October 1996.

13. A. Amin, S. Fischer, "A document skew detection method using the Hough transform," *Pattern Analysis and Applications*, vol.3, issue 3, pp.243-253, September 2000.
14. P. Y. Yin, "Skew detection and block classification of printed documents," *Image and vision computing*, vol.19, issue 8, pp.567-579, May 2001.
15. W. Postl, "Detection of linear oblique structures and skew scan in digitized documents," *8th International Conference on Pattern Recognition*, pp.687-689, 1986.
16. S. Li, Q. Shen, and J. Sun, "Skew detection using wavelet decomposition and projection profile analysis," *Pattern Recognition Letters*, vol.28, issue 5, pp.555-562, April 2007.
17. H. S. Baird, "The skew angle of printed documents," *40th Symposium of Hybrid Imaging Systems*, pp.204-208, May 1995.
18. H. S. Baird, "Anatomy of a versatile page reader," *Proceedings of the IEEE*, vol.80, no.7, pp.1059-1065, July 1992.
19. G. S. Peake, T. N. Tan, "A general algorithm for document skew angle estimation," *Proceedings of International Conference on Image Processing*, vol.2, pp.230-233, October 1997.
20. T. Akiyama, N. Hagita. "Automated entry system for printed documents," *Pattern Recognition*, vol.23, issue 11, pp.1141-1154, 1990.
21. D. S. Bloomberg, G. E. Kopec, L. Dasari, "Measuring document image skew and orientation," *Document Recognition*, vol. 2422, pp.302-316, 1995.
22. X. Zhu, X. Yin, "A new textual /non-textual classifier for document skew correction," *Proceedings of the 16th International Conference on Pattern Recognition*, pp.480-482, 2002.
23. P. Gorman, G. K. Hemantha, D. S. Guru, P. Nagabhushan, "Skew estimation of binary document images using static and dynamic thresholds useful for document image mosaicking," *Proceedings of National Workshop on IT Services and Applications*, pp.1-5, February 2003.
24. A. Hashizume, P. S. Yeh, A. Rosenfeld, "Method of detecting the orientation of aligned components," *Pattern Recognition Letters*, vol.4, issue 2, pp.125-132, May 2003.
25. P. Shivakumara, G. K. Hemantha, "Text-skew detection through contour following in document image," *Proceedings of National Workshop on Computer Vision, Graphics and Image Processing*, pp.39-44, 2002.

26. Y. Lu, C. L. Tan, "A nearest-neighbour chain based approach to skew estimation in document images," *Pattern Recognition Letters* vol.24, issue 14, pp.2315–2323, October 2003.
27. B. T. Avila, R. D. Lins, "A fast orientation and skew detection algorithm for monochromatic document images," *Proceedings of the ACM Symposium on Document Engineering*, pp.118–126, February 2005.
28. T. Akiyama, N. Hagita, "Automated entry system for printed documents," *Pattern Recognition*, vol.23, issue 11, pp.1141-1154, 1990.
29. A. Chaudhuri, S. Chaudhuri, "Robust detection of skew in document images," *IEEE Transactions on Image Processing*, vol.6, no.2, pp.344–349, February 1997.
30. J. Sauvola, M. Pietikainen, "A document management interface utilizing page decomposition and content-based compression," *Proceedings of the 13<sup>th</sup> International Conference on Pattern Recognition*, vol.3, pp.752-757, August 1996.
31. C. Sun, D. Si, "Skew and slant correction for document images using gradient direction," *Proceedings of the Fourth International Conference on Document Analysis and Recognition*, vol.1, pp.142-146, August 1997.
32. H. Ma, Z. Yu, "An enhanced skew angle estimation technique for binary document images," *Proceedings of the 5<sup>th</sup> International Conference on Document Analysis and Recognition*, pp.165-168, September 1999.
33. V. J. Traver, A. Bernardino, "A review of log-polar imaging for visual perception in robotics," *Robotics and Autonomous Systems*, vol.58, issue 4, pp.378–398, April 2010.
34. A. Alaei, U. Pal, P. Nagabhushan, F. Kimura, "A painting based technique for skew estimation of scanned documents," *International Conference on Document Analysis and Recognition*, vol.2, pp.299-303, September 2011.
35. B. V. Dhandra, V. S. Malemath, H. Mallikarjun, R. Hegadi, "Skew detection in binary image documents based on image dilation and region labeling approach," *In 18th International Conference on Pattern Recognition*, vol.2, pp.954-957, 2006.
36. L. Junjuan, T. Guoxin, "An efficient algorithm for skew correction of document image based on cyclostyle matching," *International Conference on Computer Science and Software Engineering*, vol.5, December 2008.
37. A.E. Sharif, N. Movahhedinia, "Skew estimation of persian/arabic printed documents," *Journal of Applied Science*, pp.2265-2271, 2008.

38. P. Saragiotis, N. Papamarkos, "Local skew correction in documents," *International Journal of Pattern Recognition and Artificial Intelligence*, vol.22, issue 4, June 2008.
39. T. Saba, A. Rehman, E. Boudihr, "Video motion perception using optimized Gabor filter," *International Journal of the Physical Sciences* vol.6, pp.2799-2806, June 2011.