Abstract
Scene text detection and extraction from images, where detection of text in plain background document images are not so easy, because scene text images comprises complex background, like different fonts, font size, colour, orientation etc. Hence it is an exigent research area in document analysis. We present an algorithm which involves pre-processing of images by applying wiener filter and runlength method helps to detect the text in images. This algorithm does not only detect the text in image but it also detects the blurtex.

I. INTRODUCTION
Nowadays due to increase of digital devices like cameras, mobile phones etc. Images with text content is increasing rapidly, and more interested on text present in the image convey the lot of message around the environment. A focus in this work is on image contains text information. Text is means of communication which can be implanted in any images or videos which are readable to others. There are numerous amounts of images in surrounding which contain text as information. It may be driving along the
street where we notice lots of text in images. When it comes to environment there are lots of challenges like complex backgrounds, multilingual, multi-oriented text and fonts.

1.1 Images

Images are two dimensional matrixes which are recorded and present in form of visual picture using optical devices like camera, scanner, and microscope used to capture images. These are used in many fields like entertainment, military, civil, industrial, and medical. May be any dimensional which are represented in pixel. The text images may happen to be blur while capturing the images or out of focus in natural environment. Also text detection is gaining popularity due to increase in real world application. The main focus in this project detection of text in images contains blur or non blur images. Text detection is process of detecting the text in image. Text localization means find the location of text in the image and generating bounding boxes around the text in image.

![Figure 1.1: An Image contains text.](image)

1.2 Video and Image Conversion

A sequence of video running for few seconds to minute can be converted to image i.e. Referred as frame. The video can be converted to sequence of frames where continue form of image is obtained. In this work we have taken video from 2013-2015 videos from those videos using MATLAB we have converted to frames. This video sequence can be stored in 8bit, 16bit or 24bit as frames images.

II. PROBLEM IN TEXT DETECTION AND TEXT RECOGNITION

For text recognition, binarization method is applied before sending to OCR module. Due to complex images and videos are addressed from previous method. When the image background is complex it’s difficult to exhibit binarization method. Gray-scale value can vary from one string or word distribution of gray-scale is not always bi-modal in text image. Especially in videos it may consist of two or three cluster of gray-scale value. When we apply text recognition on text it has to have capability to recognise the character in gray-scale also.

2.1 Blur Images

Due to out of focus or angle of image, motion while capturing will end up with the result blur image. If the blur image contain text which we have to detect then it becomes more difficult. A text in image is a piece of information which perceives to understand the surrounding. Due to variation in text size, fonts, style, orientation and alignment and...
also most important thing is background. The complex background may not preserve the shape of character, un-uniform spaces between characters, loss of pixels and present of noisy pixels. In order to extract the text from the blur image are difficult many research are be followed inorder to achieve the extraction of the text. But it is very difficult in natural environment.

III. LITERATURE SURVEY

1. Photo OCR: Reading Text in Uncontrolled Conditions
Commercially available OCR performs poor in text extraction [1] this approach is capable of recognizing text in a variety of challenging imaging conditions where traditional OCR systems fail. Drawback of approach it cannot classify between numbers “1” & “I” letter. They have used machine learning approaches where trained and training data is present. Here we have text input of both blurred image & regular image. Here the challenging is if lexicon is more than 50 words. There are capital and punctuation cannot be detected so.

2. A blur estimation and detection method for out-of-focus images
In this paper [2] they are estimating the blurriness in the image. Input image is blurred and out of focus images. Here the method used is canny edge detector & Hough transform. But the drawback in this method is if the background in image is blur & sharp foreground unable to detect the text in image.

3. A blind deconvolution model for scene text detection and recognition in video
In this paper [3] we have an input of blur video/image to extract the text. A Deblurring Method used blind deconvolution model that enhances the edge intensity by suppressing blurred pixels. Here they have used a uniform blur images/videos and classifying into blur and non blur images. Method to apply on non uniform blur in videos/images-Challenging.

4. Blind Image Debluring Using Dark Channel Prior
In this paper[4] input is a blur image in dark an image can contain sparse, also the blur images contain less sparse. Here challenging is Debluring image which is low intensity.
It can handle no-uniform blur in image. Complex processing techniques in kernel estimation, e.g. shock and bilateral filtering. Drawback is if the clear image doesn’t contain any dark pixels. DNoise is not handled. It fails to recognize the text in image than compared to other methods.

IV. SYSTEM ARCHITECTURE

The block diagram for text-line detection in image is proposed in Figure 4.1. In this system first image is pre-processed to normalize an illumination and remove noise and then segmented inorder to obtain a text region using run length algorithm. After the segmentation the text-line obtained from the image are extracted and aspect ratio is calculated for the region where the text is present. Abounded box is applied using a connected component.

![Block diagram of system architecture](image)

**Figure 4.1: Block diagram of system architecture.**

4.1 Image Acquisition

Text-line detection images are collected from ICDAR2013-2015 database. ICDAR provides in a bundle of datasets. Each dataset may be in format in a zip file, and contains a set of JPEG images, and XML tag of files: locations.xml, words.xml and segmentation.xml.

4.2 Pre-Processing

Pre-processing is first and foremost important step in image processing. It helps significantly in increasing the reliability of text in text detection and detecting text associated with the image. The noise in image is removed and smoothen inorder to extract text in image.
4.2.1 RGB to Gray-scale image
A colour image is converted to gray-scale. A greyscale in digital image have the value of each pixel are simple, that it have only intensity. Information stored in pixels these types of images are called black and white which also composed to gray. But the intensity varies from black and white; where black is highest intensity depending on intensity the gray is still low in intensity and white with lowest intensity.

4.2.2 Wiener Filter
Wiener Filter which reduces the noise present in an image, it is effectively filters the blur images. The better results are obtained by wiener filter for blur images. The Wiener filter reduces the mean square error across the estimated process and the obtained process. The main work of Wiener filter is to compute a statically estimation of an unknown signal using a known signal as an input and filtering that known signal to produce the estimate signal as an output.

4.2.3 Canny Edge Detector
Canny edge detector is a multistage algorithm for the edge detector, like wide edge detector. This algorithm was developed by John F-Canny in 1986. He explained the several of steps inorder to find the edges in the image. This algorithm is a good detection of many edged in image and also localization of edges are close as possible noise in image doesn’t create false edges.

4.3 Segmentation
This is an important stage where the text and no-text region is differentiated in image.

4.3.1 Horizontal Run Length
This algorithm is used in document analysis; in this algorithm the number of consecutive dark pixel is taken. When the edge is applied the black pixel counted horizontally we use the connected component. It is used to find length of the text present in image.

4.3.2 Max-Min Cluster Algorithm
The frequency of black and white pixel is present in array of CA. They form two cluster in which black and white and stores in array in CA. Now Max-Min clustering algorithm first selects the min and max frequency from the array CA. Where Max value is represent the text in image. The bounded box is applied horizontally which defines the text boundary in image.

4.4 Feature Extraction
Here we some geometric metrics like height, width and aspect ratio. The aspect ratio is calculated by the array, where the array is plotted in form of matrix. The bounded box is applied on the text region. Which is nothing but a rectangle area is applied with connected components are used to form rectangle around the text.

V. PSEUDO CODE
The pseudo code for the following modules as follows.

5.1 Main Module
Input: Text image
Output: Text-line detection
Module 1: Pseudocode as follows:
1. Begin
2. Read the input image
3. Pre-processing of image by converting to Gray-scale, applying Wiener Filter and preserving the edge by canny edge detector.
4. Segmentation using Run length algorithm, Max-Min algorithm black pixel 1 and 0 white pixel.
5. Extracting features of image bounded box using aspect ratio.
6. Text detected region extracted from non text, by false negative method.
7. End.

5.2 Pre-Processing
Input: Text image
Output: Enhanced image
An input image is colour image is converted to Gray scale (0-255). Images contain only black and white pixel it may vary according to neighbour may be gray due to vary of light intensity.

Module 2: Pseudocode as follows
1. Begin
2. Read colour image Irgb.
3. Convert Irgb to gray-scale, 0 for white and 1 for black, Ig.
4. Wiener Filter to smooth Ig obtained Iw.
5. Canny Edge to preserve the edge to Iw.
6. Extract the edge of text in image.
7. End.

5.3 Segmentation
Input: Enhanced Image
Output: text-line detected
From the enhanced image text and non-text region only text region is extracted.
1. Begin
2. Read the image from segmented Iw as F.
3. Using Run length calculates the length of text by connected component count black pixel as 1 and white 0.
4. Now Max-Min if frequency in CA selects max and min value.
5. Text is obtained by false negative.
6. Bound box height*width of connected component.
7. End.

VI. RESULTS AND FUTURE ENHANCEMENTS
The results obtained from implementation of the text detection in image detection are presented. The snapshots and graphical representation for different stages are provided. A blur input image is taken from the dataset collected from ICDAR 2013, 2015.
For Fig 6.1 Running the code for the text detection first step we convert any colour image into a grayscale image. For Fig 6.2 Gray-scale image, when applied Wiener Filter inorder to smoothen an image and reduce noise.
6.1. Future Enhancements

1) Identifying the text and non-text region in image.
2) Can be extended to text detection on handwritten documents.
3) The text with multi-orientation angle cannot be detected in this result we can also extend it for future.
4) Another filter can be added like Gaussian filter so that noise in the image gets reduced future.

VII. REFERENCES


TO CITE THIS PAPER