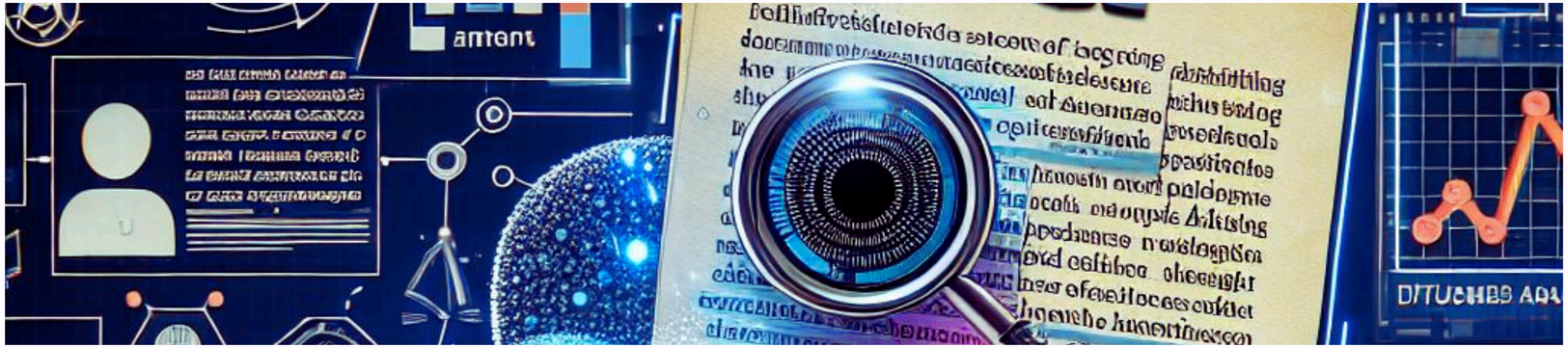
CONVERTING IMAGE TO TEXT USING PROGRAMMING

CE-712(Digital Image Processing of Remotely Sensed Data)

Group 26(CE712)

ABSTRACT



The extraction of text data from images plays a crucial role in facilitating automatic annotation, indexing, and structuring of visual information. This process, essential for applications in document analysis, digital, multimedia content management, making it a significant area of ongoing research in the fields of image processing and computer vision.

INTRODUCTION

01

Today the most information is available either on paper or in the form of photographs or videos, large information is stored in images.
The current technology is restricted to extracting text against clean backgrounds

02

Text Extraction and Recognition in Images has become a potential applications in many fields like Image Indexing, Robotics, capturing license plate information and number in traffic signals.

03

Variations Of text due to differences in Size, Style, Orientation, and alignment as well as low image contrast and complex background make the problem extremely challenging.

TEXT INFORMATION EXTRACTION (TIE) TERMS

Properties		Variants or sub-classes
Geometry	Size	Regularity in size of text
	Alignment	Horizontal/vertical
		Straight line with skew (implies vertical direction)
		Curves
		3D perspective distortion
	Inter-character distance	Aggregation of characters with uniform distance
Colour		Gray
		Colour (monochrome, polychrome)
Motion		Static
		Linear Movement
		2D rigid constrained movement
		3D rigid constrained movement
		Free Movement

STEPS IN TIE PROCESS

01 Converting colored image to grayscale

02 Binarization

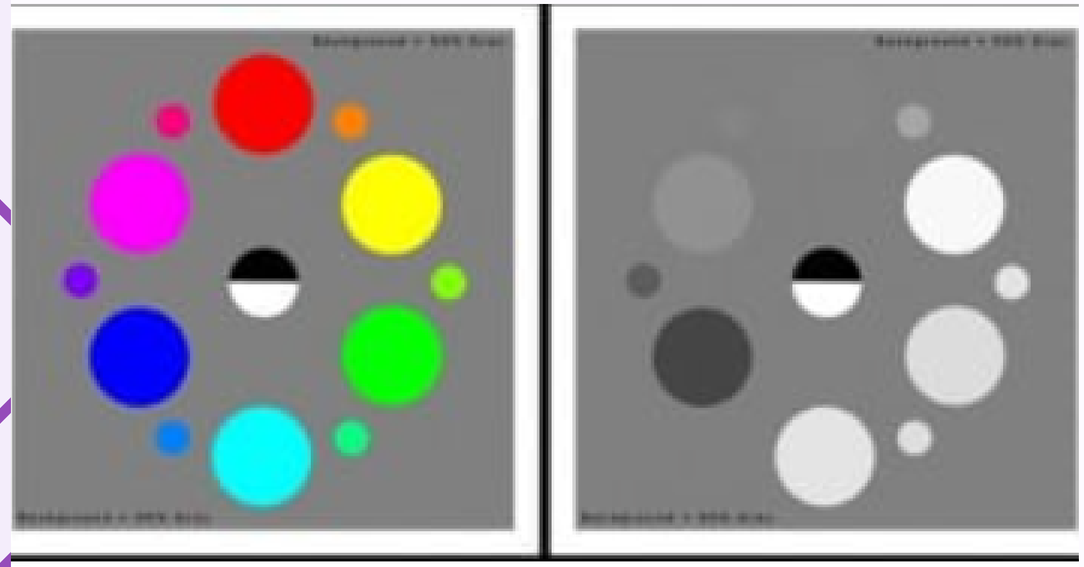
03 Connected Components

04 Horizontal and Vertical Projections

05 Reconstruction

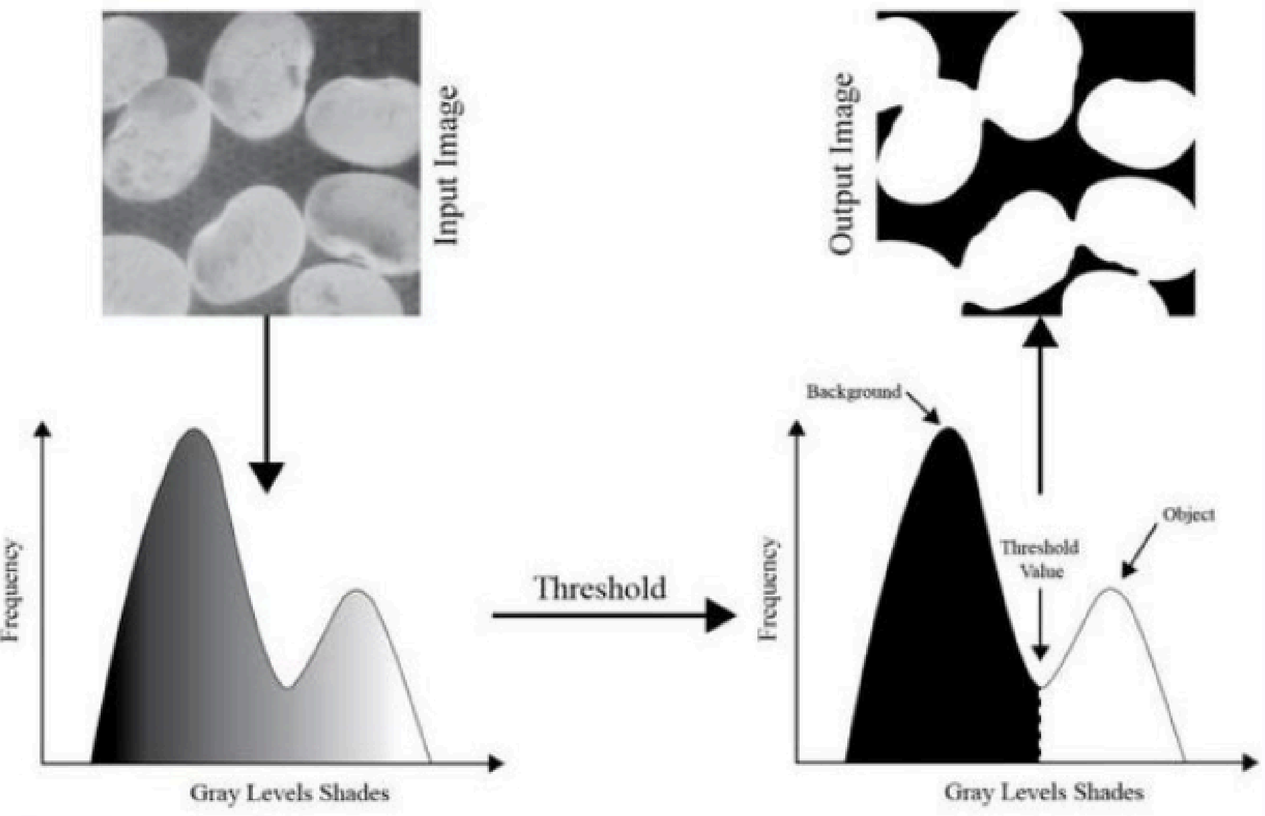
06 Recognition using OCR

CONVERTING COLORED IMAGE TO GRAYSCALE



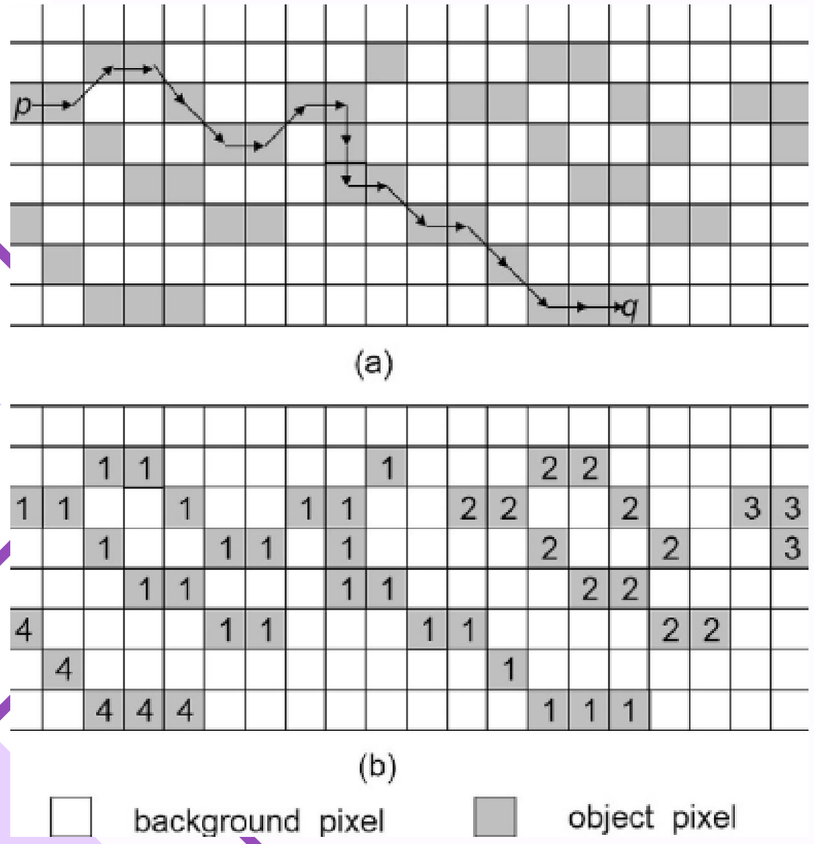
A digital color image is a color image that includes color information for each pixel. There are various models used to represent a color image for example CMY, RGB, HSI etc. Gray Scale images have range of shades of gray without apparent color. These are used as less information needs to be provided for each pixel.

BINARIZATION



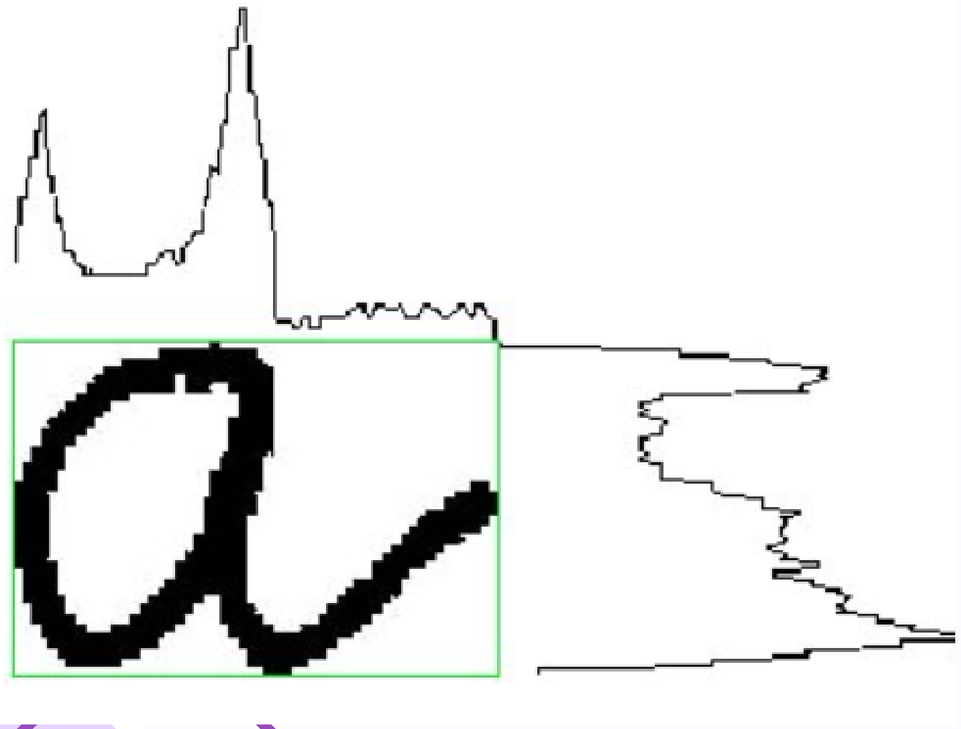
A binary image is a digital image that can have only two possible values for each pixel, 0 or 1. The name black and white is often used for this concept. To form a binary image we select a threshold intensity value. All the pixels having intensity greater than the threshold value are changes to 0(black) and pixels with intensity less are changes to 1(white).

CONNECTED COMPONENT



For two pixels to be connected they must be neighbors and their gray levels must specify a certain criterion of similarity. If P represents subset of pixels in an image, two pixels p and q are said to be connected if there exists a path between them consisting entirely of pixels in S, For any pixel p in S, the set of pixels that are connected to it in S is a called a connected component of S.

PROJECTIONS



The method is performed on binary images, it starts scanning from left side of every line and records a change in case of facing a pixel change from zero to one and again zero. Counting the change does not depend on number of pixels in this method. If the allocated amount of changes for each row is between two thresholds the row potentially will be considered text area, next we search vertically for finding the exact location of the text and ignoring these rows as a text.

IMAGE RECONSTRUCTION

Original

Result

After the extraction of text regions from images, the text regions become a bit distorted and difficult to read, thus we recover these components using the original image, the distorted and original images are compared with each other and the pixels which are erased or disfigured are recovered.

RECOGNITION USING OCR

The Optical Character Recognition (OCR) software converts each character into ASCII codes, the extracted text is stored as ASCII codes in computer memory. OCR systems have been available for a number of years and the current commercial systems can produce an extremely high recognition rate for machine-printed documents on a simple background.

TEXT IMAGE EXTRACTION

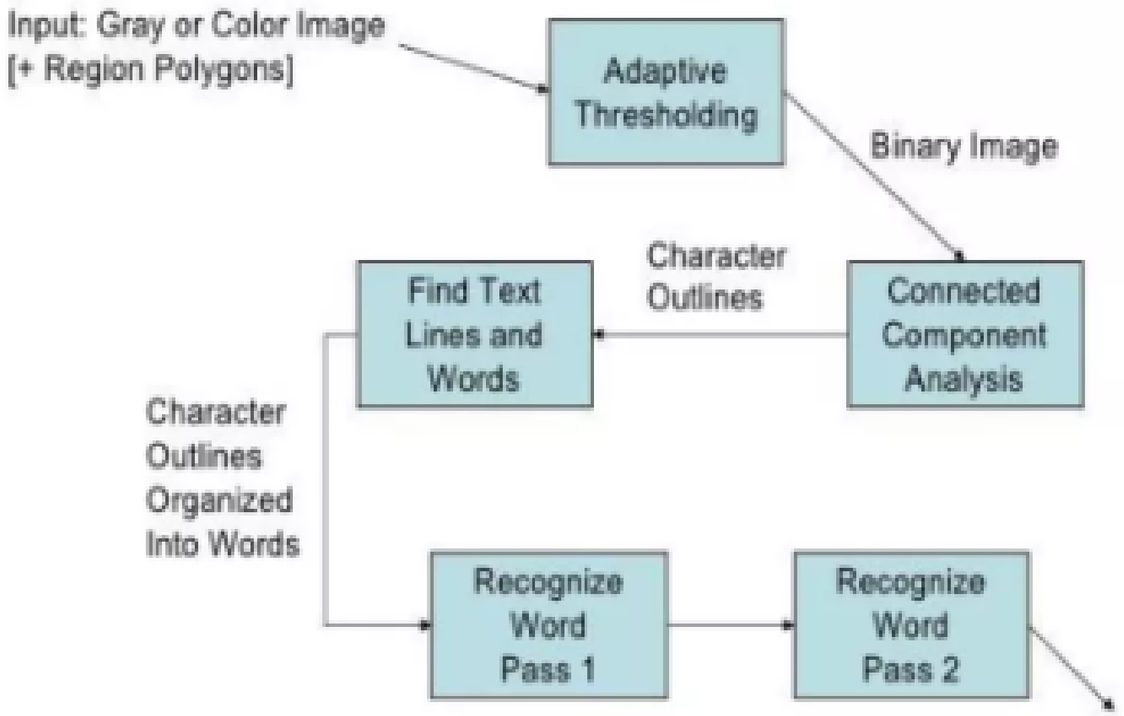


Fig. 16: Detailed Architecture of system

APPLICATIONS

CAMSCANNER

SCAN

ABBYY FineReader

Google Keep

OneNote

Google Le

PROBLEMS

01

Text on Complex Backgrounds: When text is displayed over busy or non-uniform backgrounds, accurately extracting and recognizing it becomes challenging.

02

Language and Script Variability: Recognizing multiple languages, particularly those with complex scripts, remains difficult

03

Handwritten Text Recognition: Variability in handwriting styles adds another layer of complexity, especially in forms or documents with mixed printed and handwritten tex

CONCLUSION

Even though a large number of algorithms have been proposed in the literature, no single method can provide satisfactory performance in all applications due to the large variation in character font, size, texture, color, and other attributes. With machine learning algorithms constantly being developed and improved, and massive amounts of computational power becoming readily available, OCR is rapidly advancing toward achieving higher accuracy and adaptability, bringing us closer to reliable, universal text recognition solutions across diverse environments.

THANK YOU

PRESENTED BY GROUP - 26

22B0720(ROHIT)

22B0718(CHIRAG)