DIP Project report

Sandeep Kr. Nagar 2018701015

November 26, 2018

Abstract

Aim of this project Extracting the Book and Writer's name from the cover image of the book(English) using image processing methods.

1 Introduction

In this project, the input is the photo of book cover images and the output is the text containing the text on the image including the writer name and book's name with other book information related to the book on the cover image of the book. Here I have used some methods to increase the older results with the help of some classical image processing methods.

First step is the preprocessing of the the input image like converting the colour images to gray image. After that removing the noise in the gray scale image.

Sharpening of the grey level image(high boots filtering).

Edge detection, Enhancement of Edges of text in the resultant image from earlier steps.

Binarization and Morphological operations to connecting the text's edge using erosion and dialation.

Detection- the existence or nonexistence of text in the image must be determined The final step is finding the text from the processed image using OCR(tesseract liberary).

2 Preprocessing

2.1 Denoising

Data sets collected by image sensors are generally contaminated by noise.

In the case where an image is corrupted with Gaussian noise, the wavelet shrinkage denoising has proved to be nearly optimal. There are two basic approaches to image denoising, spatial filtering methods and transform domain filtering methods. Here we have used a generalized method, Mean filtering using a 3x3 kernal.

2.2 Color to Grayscale

Here, We will convert color image to grayscale image and because of this conversion details in the image of our interest will not be affected or increase, decrease.

2.3 Sharping Image

Enhancement technique that highlights edges and fine details in an image. Which improve the results in the final image.

Geometric mean filter-: This image filtering process meant to smooth and reduce noise of an image which helps improves the edges detection. Figure-1

3 Edge detection

The local change of intensity in an image is edge. Canny edge detection algorithm and Sobel edge detection algorithm which are popularly used for image segmentation. Both of these proposed algorithms detect edges in image.Here our result is better in case of the Sobel.

3.1 Canny edge detection algorithm

The canny edge detection algorithm was proposed to enhance the edge detection process. Three important criteria were taken into consideration for this purpose.

Figure 1: output image after applying the sharpnin

The first and most important criterion was to detect all the important edges in the source image.

3.2 Sobel edge detection algo.

The gradient based operators detect thick and rough which did not give appropriate results in further matching process

The main advantage of Sobel operator is its simplicity which is because of the approximate gradient calculation.

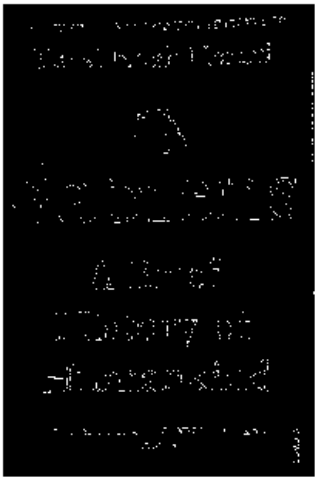
Figure-2

4 Binarization

Binarization helps in recognising text and symbols.

4.1 Adaptive Mean thresholding.

The conventional thresholding techniques use a global threshold for all pixels, whereas adaptive thresholding changes the threshold value dynamically over the



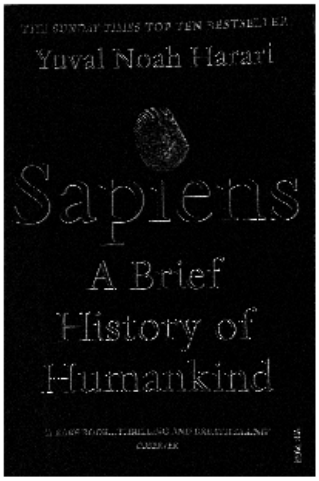


Figure 2: Results for the Canny and Sobal.

image.

Adaptive thresholding techniques to choose the accurate method for binarizing an image based on the contrast, texture, resolution etc. of an image

4.2 Adaptive Thresholding with gaussian weighted average

This adaptive threshold method by using the concep of the FIM.

The algorithm utilizes statistical quantities such as mean, standard deviation, and variance to define a new adaptive and automatic threshold based on two-frame and three-frame differencing.

4.3 Fixed thresholding

The simplest thresholding methods replace each pixel in an image with a black pixel if the image intensity is less than some fixed constant T (that is,), or a white pixel if the image intensity is greater than that constant.

Figure-3

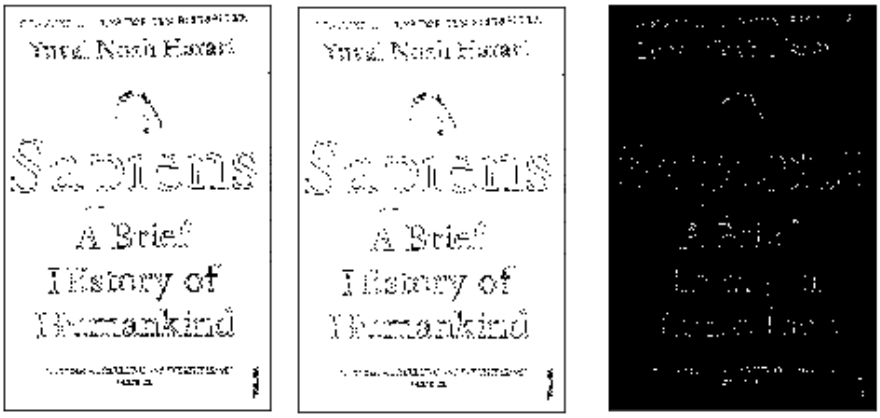


Figure 3: Thresholding with Adaptive mean,Adaptive Gaussian and Fixed

5 Morphological operations

Image after the thresholding, some edges of the text in the image are not connected because there is no perfect thresholding which can works for edge detection.To connect these edge and improving the text information in the image, we perform some morphological operations.

Dialation and Erosion. Figure-4

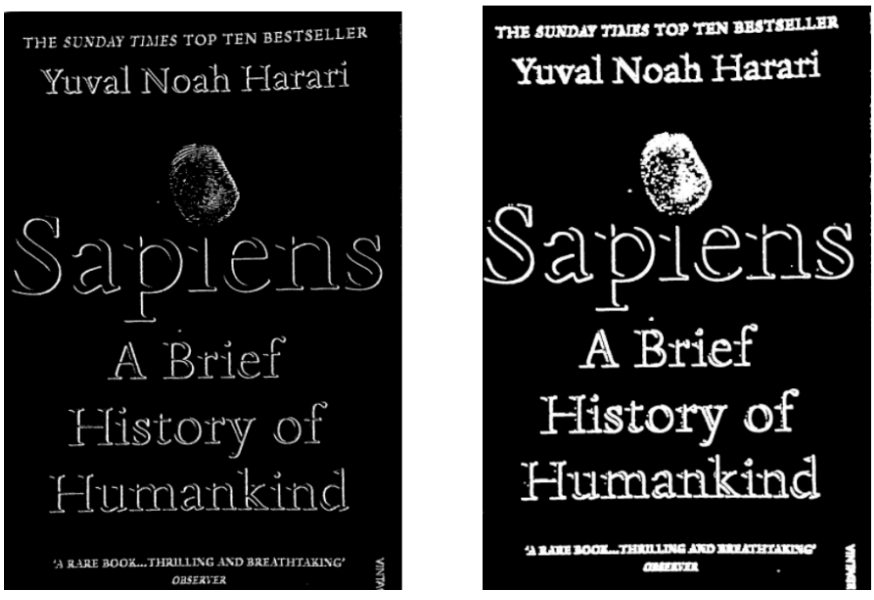
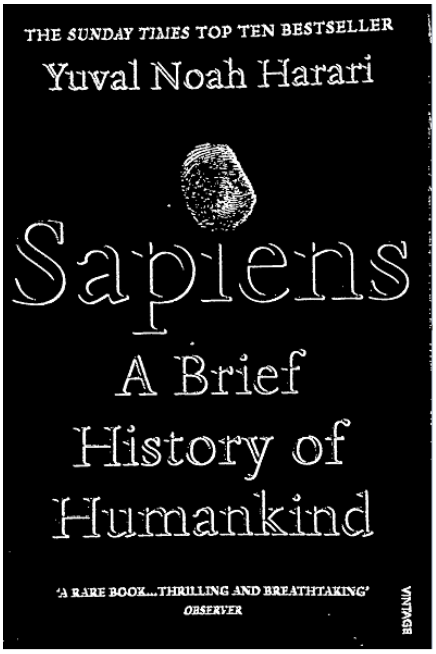


Figure 4: Images befor and after the morphological oerations

6 Optical Character Recognition

To extract the text from the image here we used the Tesseract python library. If we give the orignal image as input and the image after the complete processing we proposed here in this project, there is much more difference in between the resultant of the both.

Figure-5



Input Image

```
THE SUNDAY TIMES TOP TEN BESTSELLER

Yuval Noah Harafl

1%,.-
zit-1:.

- -. -_--

. ·_ ____··

.j/I. .1;

1-,- ‘_--.~,_ 731.. ._

:5: ""'.L-'_' R I _

_- 3:. .;--.a-_ .


A Busf
Higmry of
Hum. ankin d

1-! RARE BOOK...THR1]_LING AND BRLAH-ITAKING’
OBSHH-R

HBVJNM
```

Output Text

Figure 5: Final Text from the result from processed image.

6.1 Other Results from different

Figure-6

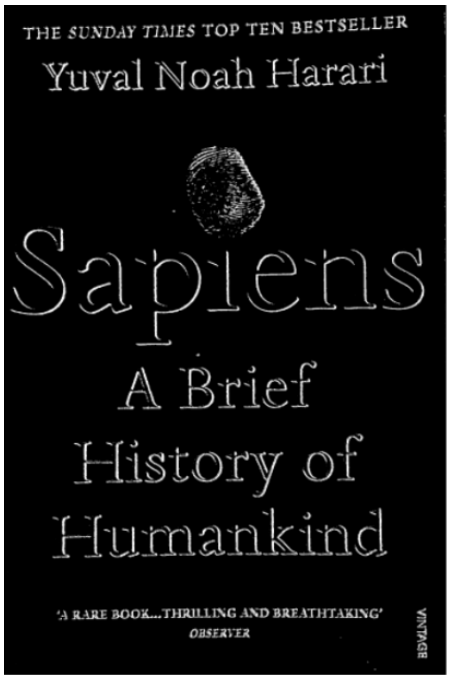


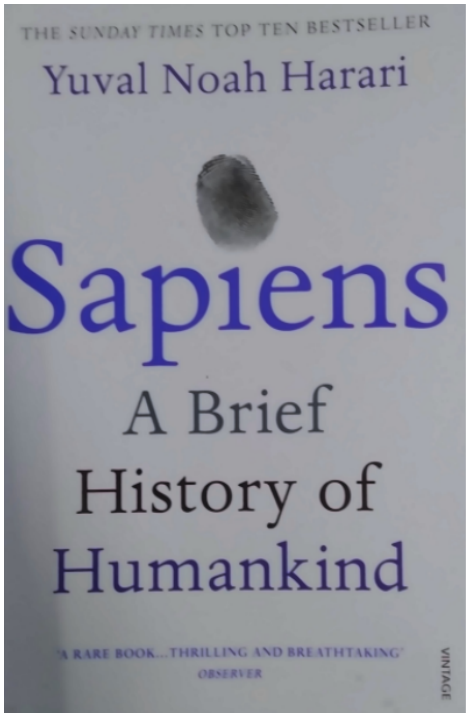
Image without morphing oeration

```
THE SUNDAY mms TOP TEN BESTSELLER g!

Yuval Noah Harafl


E115 rory Of
Hum ankin Cl,

1-! RARE BOOK...THR1]JJNG AND BRL-i'I'J-ITAKING'
OBSH'HER

HBVJNM
```

text Output

Figure 6: Final Text from the Image without morphological operation.

Figure-7




```
/n THE Si VD.” “Ml-‘5 TOP TF\ BESTSELLE-
Yuval Noah Harari


C
S ap1ens
A Brief
w History of



1

0| EEEEEE
```

Figure 7: Final Text from the result from processed image

6.2 Result comparison for Different cases

When we give input image without any processing then in the result only dark colour text is extracted, in the Tesseract.

If we gives the input image with the processed then the result is more better then the all other cases.

6.3 Failure cases

Very less accuracy when book cover is more contrast, reflective. When the text colour and the background colour are same. Tesseract will fail with noisy background. Sobel is less sensitive to noise but computationally it is slower. To connect the edges we also can use Hough Transformation. below some image are shown for which effecency is low.

Figure-8

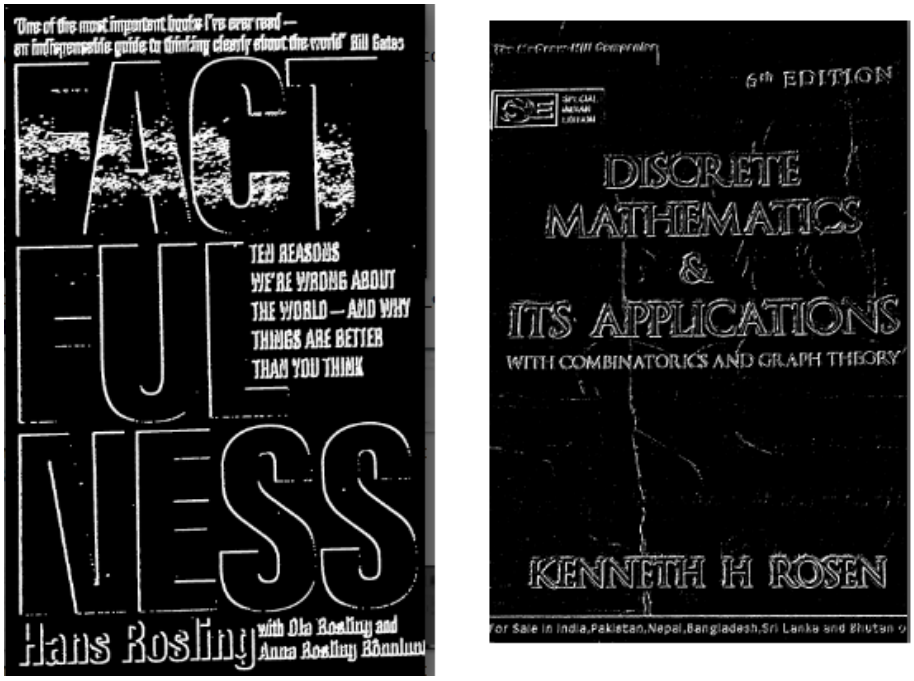


Figure 8: Final Text from the result from processed image

7 References

1. Christian Wolf, Jean-Michel Jolion, Françoise Chassaing:Text Localization, Enhancement and Binarization in Multimedia Documents. ICPR (2) 2002: 1037-1040

2. Image Binarization using Otsu Thresholding Algorithm . Jamileh Youse

3. LOCATING TEXT IN COMPLEX COLOR IMAGES YU ZHONG, KALLE KARU and ANIL K. JAIN

4. TEXT DETECTION AND CHARACTER RECOGNITION USING FUZZY IMAGE PROCESSING Mohanad Alata — Mohammad Al-Shabi