Document scanner with image enhancement and compression

Project report for IITB EE610 Image Processing 2021

Lovepreet Singh
Electrical Engineering (EE1)
Indian Institute Of Technology
Bombay, India
213070020@iitb.ac.in

Surbhika Rastogi
Electrical Engineering (EE5)
Indian Institute Of Technology
Bombay, India
213070056@iitb.ac.in

***Abstract*— Paper documents are frequently scanned to images for quick e-sharing. In such cases, mobile scanner applications are quite handful which improves upon the irregularities in mobile-clicked pictured and convert them to more-readable formats. This project aims at handling challenges of a document scanner application e.g., Adobe Scan, Cam Scanner etc. The main features achieved through this course project are: i) automatic recognition of the document from the captured image which may include background also, ii) Improve the perspective view of the document and reshape it into rectangular form, iii) color enhance the image to make it more readable, iv) option to remove possible blurring of the document due to camera motion, v) compress/resize the image to desired size, and vi) convert the selected images to pdf format for sharing purpose. Several spatial and frequency domain transformation techniques are applied to achieve the aforementioned tasks. Morphological operations are applied to improve readability of the content and wiener filtering method to remove the motion blurring. The aforementioned tasks achieved through this project are able to convert a mobile phone clicked image to a sharable document. The motion de-blurring using wiener filtering slightly improves the readability but is not sufficient to completely eradicate the issue. Some more advanced methods involving machine learning techniques are required to be able to reliably predict the blurred characters in the document.**

I. INTRODUCTION

Nowadays, document scanner becomes more popular to scan images and making consolidated PDF. All these scanner try to mimic original scanner. While taking photograph of any A4 size paper the mainly 3 problems arise :

1) Image contain background which is not useful
2) Consolidate images in the form of PDF
3) There should be good contrast between paper back ground and writing color
4) Compress image
5) Remove motion blurring

Our project deals with these 5 problems. We have used some basic image processing techniques such as wiener filtering, variable thresholding, edge detection to solve these problems

The problem such as motion de-blurring, cropping to exact shape and background enhancement is difficult to solve. We have found contours of different shapes and select proper contour according to the image shape. According to contour (usual Trapezoid) we have crop the image and remove background.

Sometimes camera gets into planar motion and image gets blurred. We have motion deblurring discussed in [1] to recover text to a certain extent.

We have used variable thresholding using moving average to enhance the image and to give scanner like feel.

II. BACKGROUND AND PRIOR WORK

Motion deblurring: - Motion deblurring can be achieved through wiener filtering.
To achieve that, we first need to estimate degradation function. [1] clearly gives derivation of degradation function of motion blurring.
We are showing result here

$$H(u,v)=\frac{T}{\pi(ua+bv)}\times\sin(\pi(ua+bv))\exp(-j\pi(ua+bv))$$

H(u,v) is degradation function in Fourier domain.

u and v are frequency variable

a – motion in x direction

b – motion in y direction

T – duration of exposure.

T, a and b are hyper parameters and will set according to human perception

The recovered image is given as

$$F(u,v)=\frac{|H(u,v)|^2\,G(u,v)}{H(u,v)(|H(u,v)|^2+K)}$$

K is hyperparameter

G(u,v) –degraded image

Image enhancement:- we are using variable thresholding using moving average. [1] provides details about it. Moving average at k+1 step

$$m(k+1)=m(k)+\frac{(z_{k+1}-z_{k-n})}{n}\qquad \text{for } k>=n+1$$

n – no. of points used in computing the average. n is usually 5 times the average stroke width

z_{k-1} – intensity of point encountering in the scanning sequence at step k+1

thresholded image is given as

$$g(x,y)= \begin{cases} 1 & \text{if } f(x,y) > bm_{xy} \\ 0 & \text{if } f(x,y) <= bm_{xy} \end{cases}$$

m_{xy} = moving average at point (x,y) in the input image

b = hyperparameter

we have used morphological operations(opening, closing, convex hull) in image cropping. Lets discuss its background

Opening : -Opening generally smoothes the contour of an object. The opening of set A by structuring element B,

$$A \circ B = (A \ominus B) \oplus B$$

Closing:- Closing also tends to smooth sections of contours but, as opposed to opening, it generally fuses narrow breaks and long thin gulfs, eliminates small holes, and fills gaps in the contour. Closing of set A by structuring element B,

$$A \bullet B = (A \oplus B) \ominus B$$

III. METHODOLOGY

A. Document Recognition and cropping algorithm: -

Following steps were performed to be able to recognize the image area covered by the document:

1) Blur the image slightly so the contents written over it fades away
2) Apply thresholding to separate the paper from background
3) Perform morphological operations to get rid of extr a/tiny areas falsely detected
4) Find the possible closed contours in the thresholde d image
5) Approximate the contours to polygons and separate the ones which have 4 sides (since documents are usually considered to be paper with 4 sides)
6) Find the area of all contours approximated with 4 si des
7) Select the polygon with largest contour area to be t he document area

Following steps were performed to warp the recognized document area to rectangular shape:

1) Find approximate height and width of the area recognized
2) Map the corner coordinates of selected polygon to t he resized image
3) Warp the polygon to resized image

B. Algorithm used for motion deblurring:-

1) Compute Fourier Transform of degraded image
2) Select appropriate values of all hyper parameter
3) Compute H(u,v) and magnitude of H(u,v)
4) Put all these values in the expression of F(u,v)
5) Repeat till all pixel put pixel in dummy image
6) Take IDFT of dummy image

C. Algorithm for image enhancement :-

1) We have to go in zigzag pattern
2) First we will traverse row wise; then change to the next column after iterating each pixel in single row

3) Calculate moving average and temporarily store in array
4) When average of one row calculated then it is stored in another array and temporary array is reset
5) Select appropriate value of coefficient b
6) threshold the image

D. Algorithm for image compression:-

An image can easily be resized to a given no. of pixels by interpolation methods. But the challenge is due to the fact that there is absolutely no way to estimate the size of the resized image after compression standards (such as .jpeg or .png) used to store the image on disk or for sharing. Usually the disk space matters when we need to limit the size of a document or share through email/submit on form. This is because the compression factor of these standards depends upon what is exactly stored in the image. It means a completely blank image will take lesser disk-space even if it has equal number of pixels compared to the image which may have stored a real image of an animal.

So iterative method is used to measure and resize the image to finally downscale the size upto the maximum limit provided by the user of application with some permissible error in size.

E. Algorithm for conversion to pdf:-

Main challenge while automatic conversion of images to pdf is that we need to resize all the images such that they fit in the pdf area.
So firstly, we first resize all the images to a given factor. Then keep the aspect ratio of individual images constant to avoid undesired stretching. Finally used FPDF library function to add new pages and convert to pdf format.

IV. EXPERIMENT AND RESULT:-

Following image is selected as input image for document recognition.

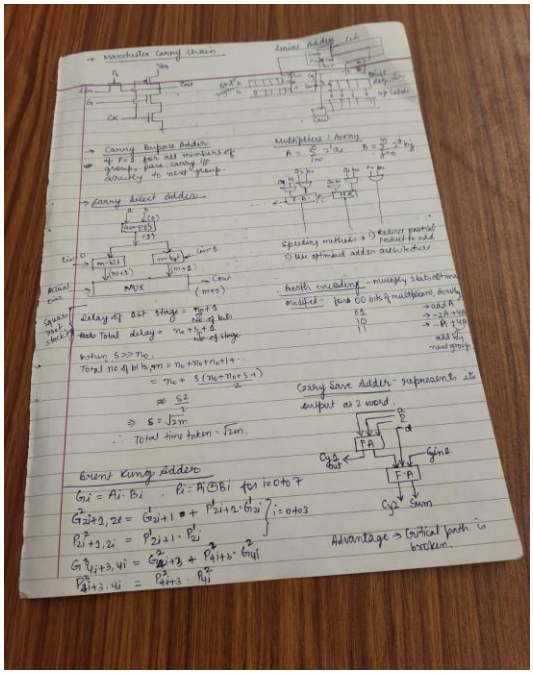

Figure 1

The above image has multiple challenges:

1) *The background (table pattern) has also some edges*
2) *It is perspectively taken. It means it's not proper rectangle or square.*

The output of document recognition algorithm detects multiple contours shown below.

.

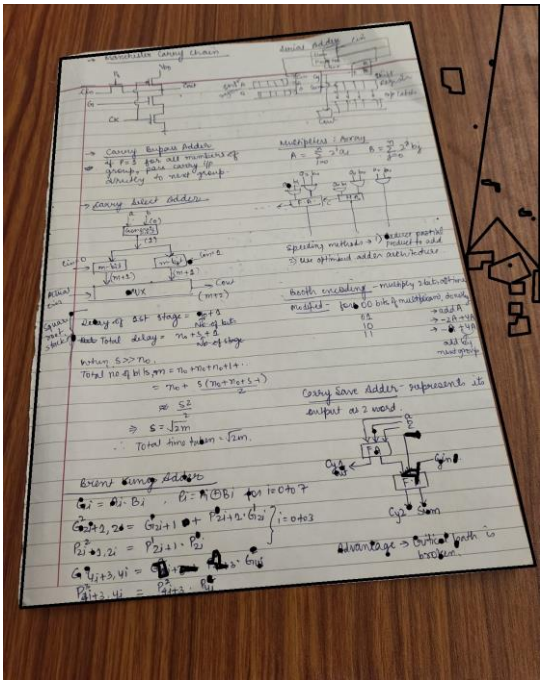

Figure 2

The above image shows that there are several false contours also. The contour with 4 sides and largest area is selected and further prospectively warped to get the final image of document which has to be further enhanced.

The document image achieved after warping is as follows:

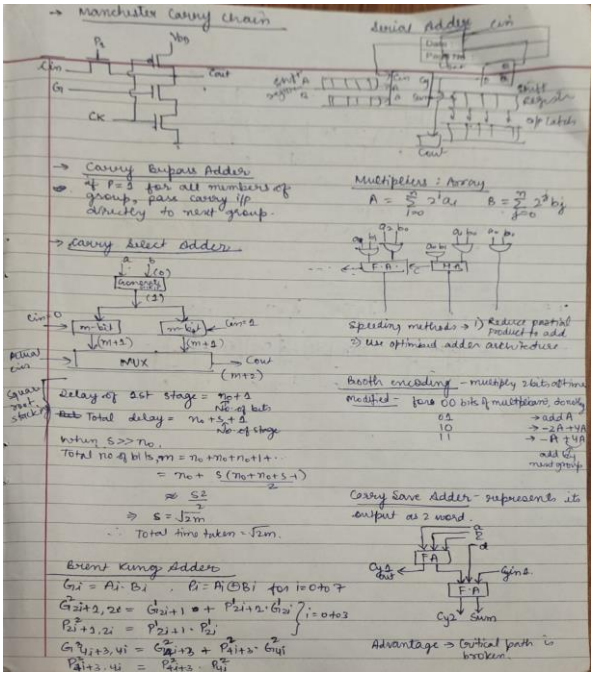

Figure 3

In another image example, we show the variable thresholding results.

Input image:

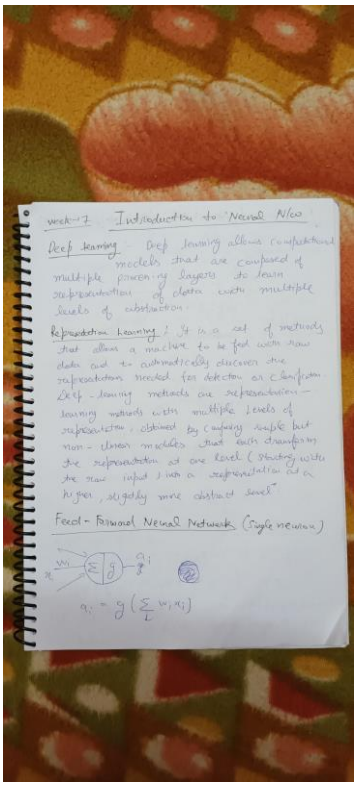

Figure 4

Output image after automatic recognition and cropping:-

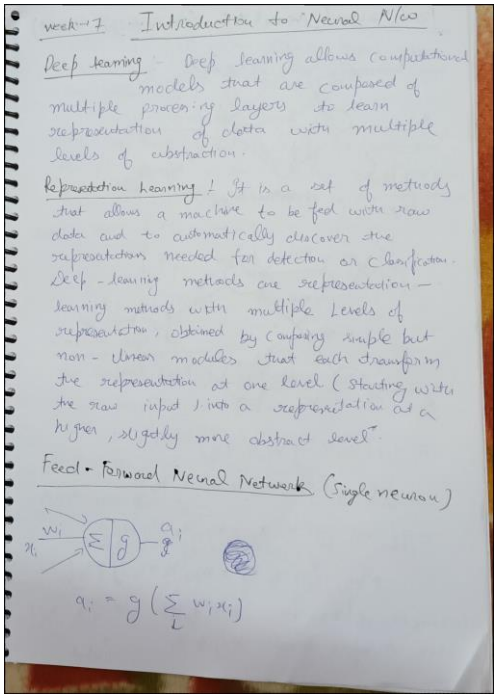

Figure 5

Output of image enhancer (Magic filter):-

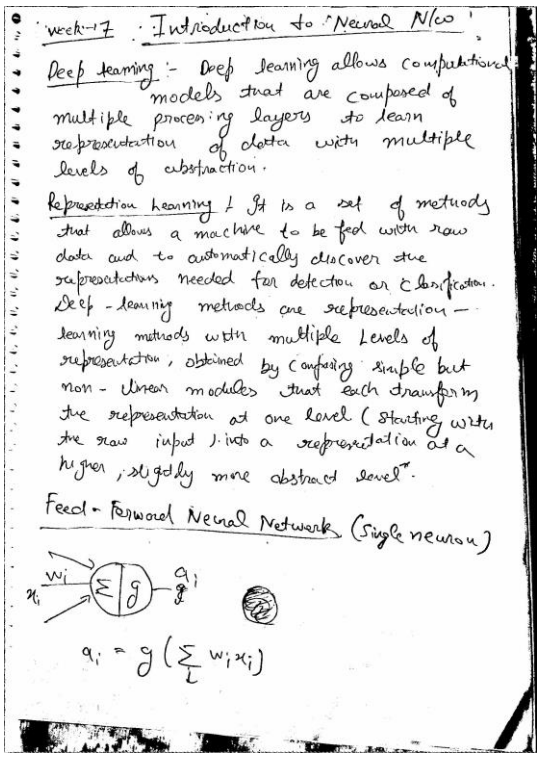

Figure 6

In third example, results of motion de-blurring feature are demonstrated:

Input image to wiener filter:-

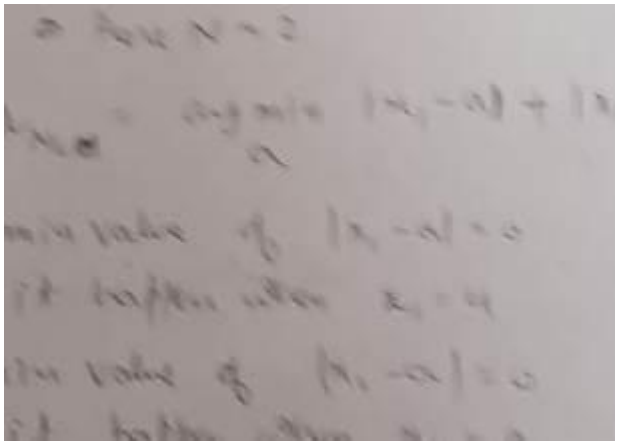

Figure 7

Output of wiener filter (small improvement):

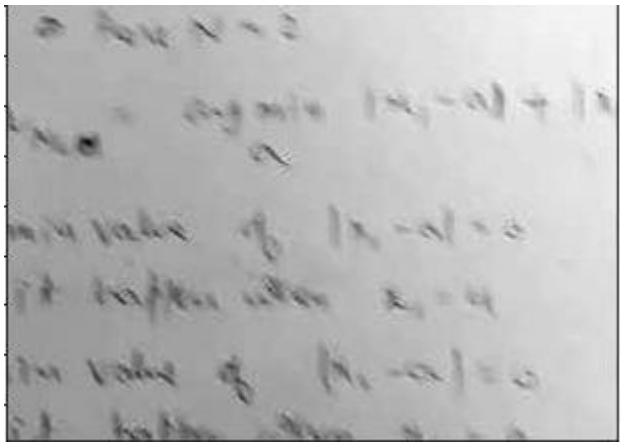

Figure 8

Image compression feature is demonstrated through following example.

Target file size: 2kB, Delta: 0.01

Input image and size:

Figure 9

Compressed image:-

Figure 10

V. Conclusion and Future Work

This overall project is great learning for our team. We have implemented some algorithms which are studied in coursework in theoretical form. Algorithm such as wiener filtering, variable thresholding and morphological operation are studied earlier in the coursework.

In future one can implement algorithm such as all the hyper parameter are selected automatically instead of choosing manually. Good compression ratio can be achieved using DCT or WHT block encoding algorithm, which are used in JPEG images. Also variable thresholding gives breaks in written work, it can be removed through morphological operations

Contribution of Team Members:

The project was divided into following 6 tasks:

1. *Automatic recognition of the document from the image and crop it*
2. *Resize the document to rectangular shape from usual trapezoidal shapes usually received when clicking a picture of paper due to perspective angle.*
3. *Enhance the document by applying transformations such as grey scale, contrast etc.*
4. *Option to remove motion blurring in images due to moving camera*
5. *Image compression to a specified reduced size*
6. *Convert the set of images to compressed pdf format.*

Tasks 1,2, 5,6 – completed by Surbhika Rastogi

Tasks 3,4 – completed by Lovepreet Singh

References

[1] Digital.Image.Processing.4th.Edition.by Gonzalez woods

[2] https://www.youtube.com/watch?v=Wka_XhcZAcQ

[3] https://www.geeksforgeeks.org/python-pil-image-resize-method/

[4]https://stackoverflow.com/questions/46689428/convert-np-array-of-type-float64-to-type-uint8-scaling-values

[5]https://docs.opencv.org/4.x/d1/de0/tutorial_py_feature_homography.html