DIGITAL IMAGE PROCESSING

PROJECT REPORT

"DIGITALIZATION OF BILLS"

Biller	Due Date	Amount
StarHub 1.12042314	30 Nov,2010	78.3
StarHub 1.13638826	26 Nov,2010	135.93
Sing Power 8918424485	10 Nov,2010	186.66
SingTel 11487286	09 Nov,2010	134.09
StarHub 1.12042314	31 Oct,2010	83.98
StarHub 1.13638826	29 Oct,2010	135.93
Sing Power 8918424485	12 Oct,2010	202.51
SingTel 11487286	11 Oct,2010	127.86
StarHub 1.13638826	06 Oct,2010	131.71

Guided By

Dr. Vineet Gandhi

Team Buildozers

Manan Bhandari(201431166)

Shivang Nagaria (201401146)

Toshad Salweker (201401167)

ABSTRACT

Image processing has been taking a larger space in the research areas. Every research is, in some or the other way, related to image processing.Improvement of pictorial information for betterment of human perception like deblurring, de-noising in several fields such as satellite imaging, medical imaging etc are renewed research thrust. Specifically we would like to elaborate our experience on the significance of computer vision as one of the domains where physical objects or information is processed and stored as piecewise bit information in the computer. Digitising bills can be done using MSER and OCR (Optical Character Recognition) method. OCR engines have been developed into many kinds of object-oriented OCR applications, such as receipt OCR, invoice OCR, check OCR, legal billing document OCR. Digital image processing allows the use of much more complex algorithms, and hence, can offer both more sophisticated performance at simple tasks, and the implementation of methods which would be impossible by analog means.

INTRODUCTION

Converting physical information into digital data is always been a goal for technology. Storage and management of digital data is always easier than storing a bunch of hard copies in a long run.

People try hard to keep physical information organized, but it is difficult. Bill management is one of the problem they struggle in their day-to-day life. Our approach solve this problem by using different techniques of Digital Image processing.

Digital image processing is the use of computer algorithms to perform image processing on digital images. As a subcategory or field of digital signal processing, digital image processing has many advantages.

Our solution takes the image of a bill, extracts useful information and stores it in digital form. Furthermore, it works for different kinds of bill, whether it is printed or handwritten. We wanted the our solution to work for all kinds of bills, shop bill, computer generated receipts, etc.

APPROACH

Our software will extract useful information from the receipts like shop name, customer's name, items purchased and their cost and total amount and save it in digital form. But How?

PRE-PROCESSING

The receipt image should be pre-processed for better results. Needs to handle issues like:

- Removing noise and degradation from the image
- Cropping of other objects and getting the bill as the image with correct straight alignment
- Image alignment

Cropping out other objects

The receipt image could contain other unwanted objects too. Here, we make use of the fact that the receipt has a white background and that it is written on a rectangular sheet and located near the centre of the image.

Here, we use quite a few things. First we binarize the image using a thresholding algorithm which considers both local and global mean, and uses a high threshold value. Through this, we expect to get the white paper of the receipt as 1s and the rest as 0s.

Next, we find the edges in this image, and then try to find the principal lines using hough transform. We will get the sides of the receipt. Now we close the

image using ‘line’ structuring elements parallel to the edges of the receipt and of a significant length. This step will remove any text from the receipt. Finally, we open and close with a large disc structuring element to remove any unwanted artifacts.Thus, we will get a binary image where the white part corresponds to the area of the receipt in the original image. We can use this as a mask and apply it on original image to get the relevant part.

Image Alignment

As our solution needs image of a bill, we need to first align the bill accordingly so that text detection can be improved. We’ve used Hough Transform for this purpose.

A bill always have some vertical and horizontal lines which may corresponds to its sides or the table which contains the list of items purchased. In case, if the image is not aligned correctly then these lines won’t be horizontal or vertical.

First we rotate image by some angle (-5° to +5°), it is observed that performance increased after including this step. It helps to detect principal lines in the images which are just slightly deviated from the vertical.

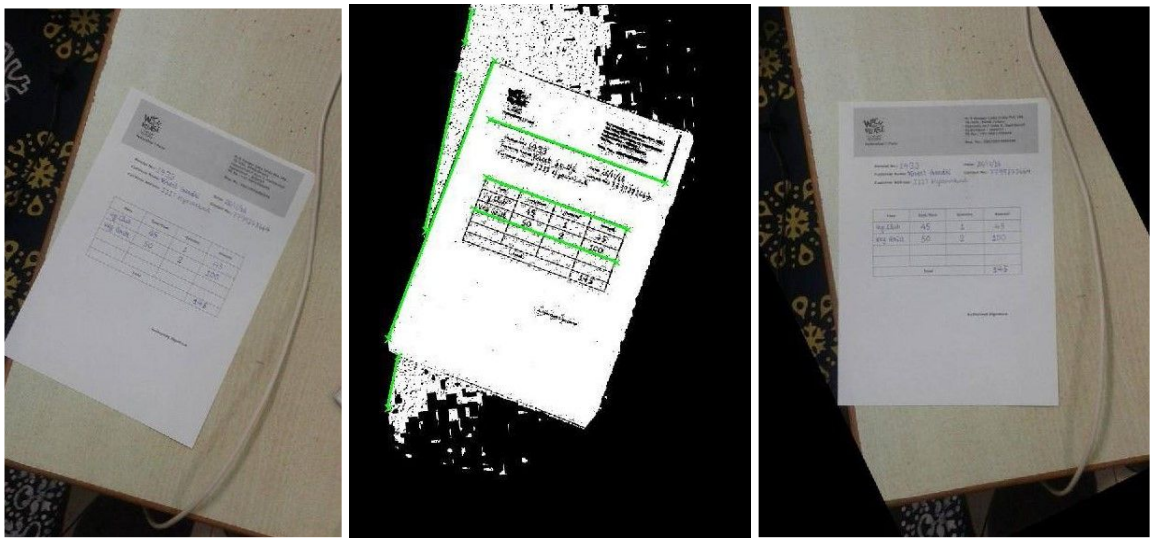
Using Hough Transform, we find principal lines in the image which mostly will be overlapping with the bill’s sides and the table. The angle from which these lines are deviated is the angle (let’s say this angle, A), we need to rotate the image in opposite direction to get aligned image. We have a list which contains the angles from which the principal lines are deviated.

There are three ways to get A:

I. Average of all these deviation angles.
II. Minimum of all these deviation angles.
III. Maximum of all these deviation angles.

It is observed that the third option performs better than other options. We also need to filter out some principal lines which are deviated too much because it is even possible that some of the lines are not horizontal or vertical in first place.

After getting angle A, we can rotate image in opposite direction to get aligned image. Now, this image is ready for next step.



PROCESSING

Recognizing text in images is useful in many computer vision applications such as image search, document analysis, and robot navigation. For detecting the letters and hence the words written on the bill, be it printed or handwritten, we have used OCR method.

MSER (Maximally Stable Extremal Regions)

It proceeds by first sorting the pixels by intensity. This would take O(n) time, using binary sort. After sorting, pixels are marked in the image, and the list of growing and merging connected components and their areas is maintained. In practice these steps are very fast. During this process, the area of each connected component as a function of intensity is stored producing a data structure. A merge of two components is viewed as termination of existence of the smaller component and an insertion of all pixels of the smaller component into the larger one. In the extremal regions, the 'maximally stable' ones are those corresponding to thresholds where the relative area change as a function of relative change of threshold is at a local minimum, i.e. the MSER are the parts of the image where local binarization is stable over a large range of thresholds.

So in our project, we have used MSER on text to recognize each character.

Optical Character Recognition (OCR)

The OCR function provides an easy way to add text recognition functionality to a wide range of applications. OCR works best when there is uniform background. If there is a non uniform background, pre-processing was done to get better results.

When this method is applied, we get each character and its bounding box which has its coordinates. An overlap ratio has been set which decides which all characters to combine on the basis of bounding boxes to built all the words in the image. Hence, we obtain all the words and hence its bounding boxes with the coordinates.

The approach we used next was pretty trivial. Every bill has the basic information asked like name, bill number, date, phone number etc. and the important information about the products bought like product name, amount, price and total amount of the bill.

For obtaining the basic information

When all the words are obtained, we search for the words like “Name”, “Date” , “Bill No.” and corresponding to their coordinates we look for the words written just after them from the coordinates we have and get the data.

For obtaining the important information

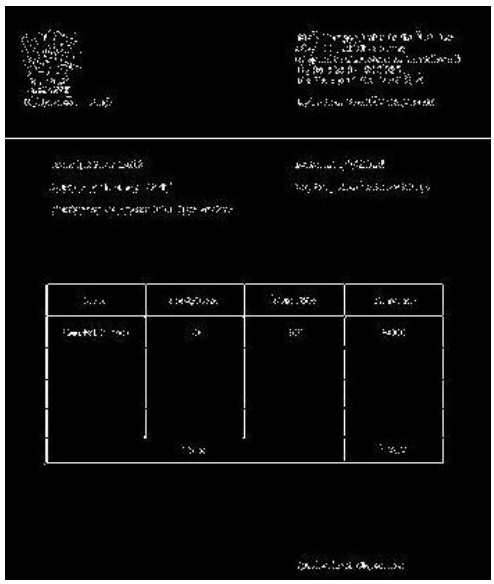
Generally all the information like product name, quantity, amount etc are written in 2 ways on the bill : either in tabulated form or normal form with some space between product and their cost.

a) **For the tabulated form**

   Using Hough transform that was used earlier for image alignment we can find lines that makes that table. We filter out the result on the basis of length of the lines and their position.

   With the help of the coordinates of the bounding boxes of each word obtained from OCR method earlier, we can find all the words and numbers within the table and arrange them according to the category they belong.

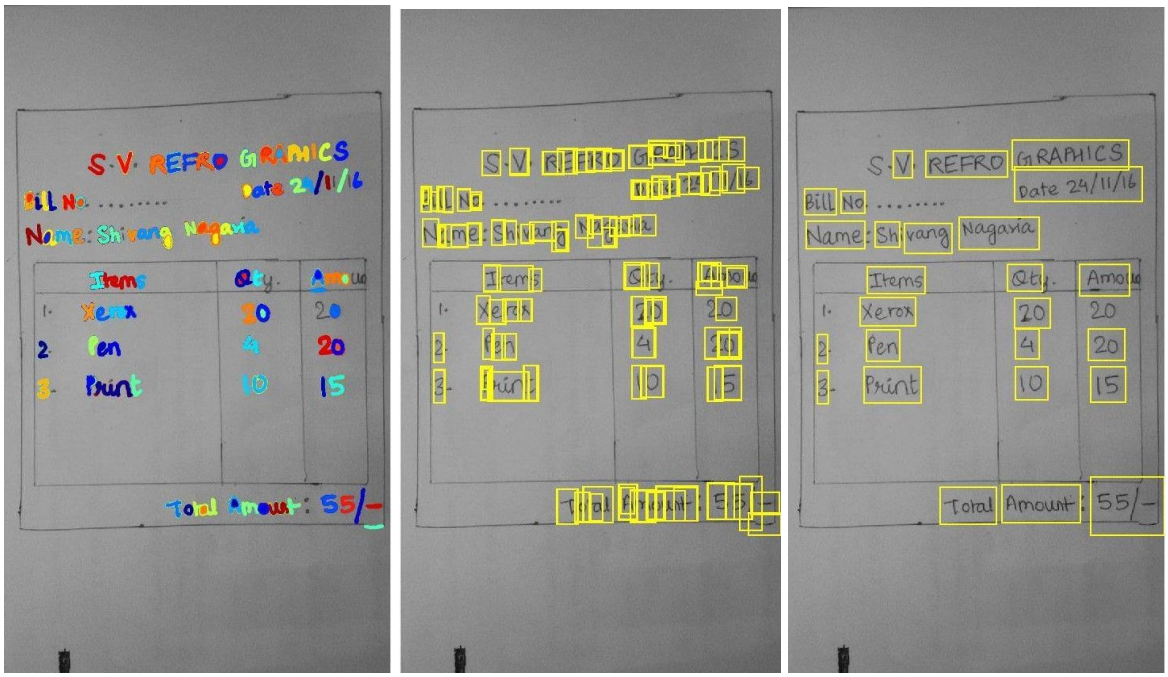
It is also possible that some principal lines are too close to each other, in that case average value is used for sending words to appropriate category.



b) For the other bills

There is always some gap between the product name and the amount or quantity bought. As the image we are dealing with hass already been converted to black and white image, we can find the white patch that exists between those columns. This can be done by opening the image with a rectangular patch of a specific ratio as that of the image. On the

left of that will be the product names which can be obtained by comparing the pixels from bounding box. And similarly for the right patch to get the amount of each product and hence the total amount.

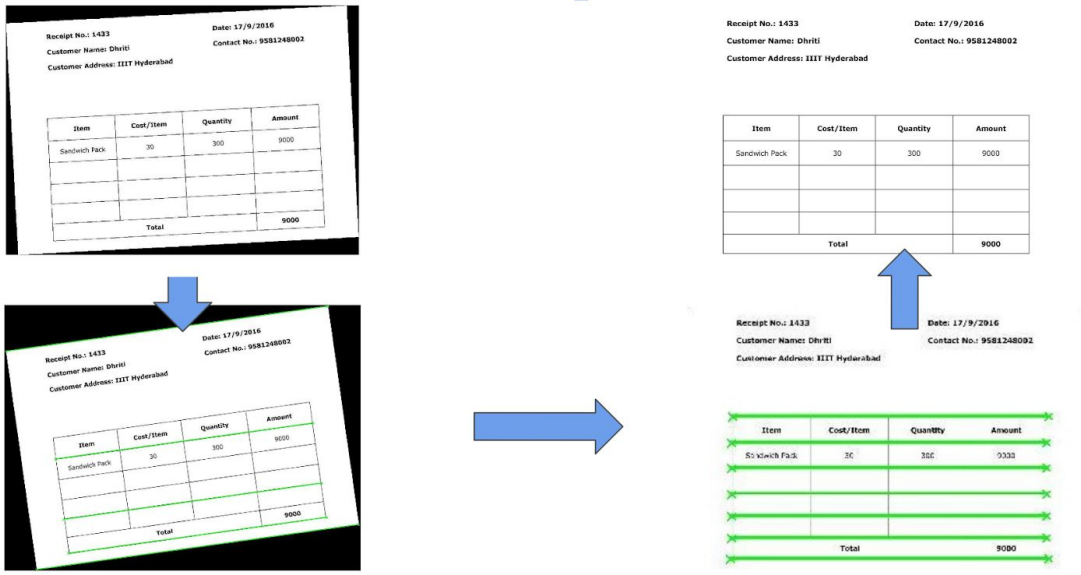


POST-PROCESSING

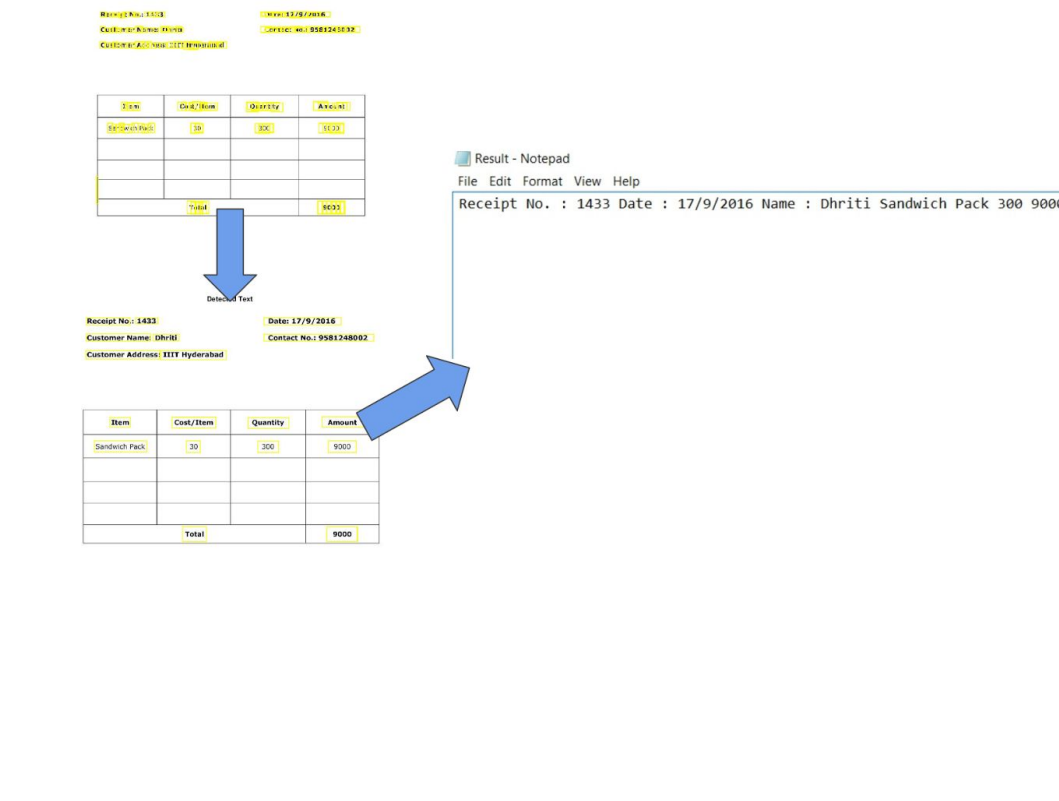
As the information is obtained when the code is running, it is simultaneously being saved in a text file. By this, we get the basic as well as the important information.

RESULTS

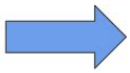
Example

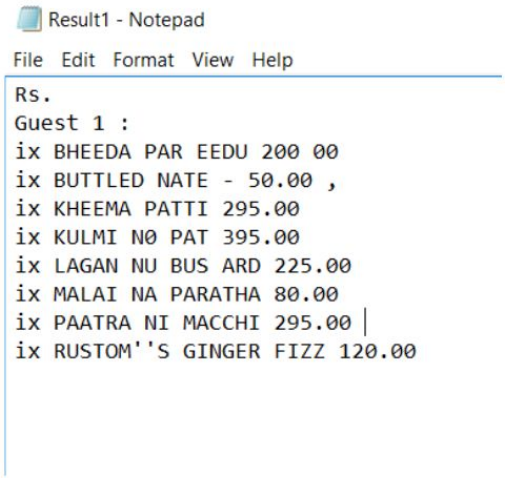


Example



Example 2





FURTHER IMPROVEMENTS

1. Number of true principle lines detected can be increased if we also consider their length. If a line's length is larger than some threshold (let's say ⅛ of the length of image) then it can be consider as a true principle line.
2. Adaptive thresholding can be used in place of local thresholding for better results.
3. For the handwritten bills, we can use template matching for each character and number, and then use bounding boxes to club the letters to form the words. This would definitely give better results than the technique used in this report.

ASSUMPTIONS

1. Bills can be two kinds, one which list items in a table and another without table.
2. Each bill will contain keywords like Receipt No. , Name, Date before relevant fields. Bills may have Invoice No. instead of Receipt No. or customer's name is written below 'Name', not beside it but we've

ignored such cases for sake of simplicity.
3. Table which contains lists of item contains ‘Item’ as column header.
4. The line where last row of the table ends should be the last line of the image and it should be horizontal.

CONCLUSION

In this project, we addressed a problem for which not much solutions are available. Using this solution, bills can be stored and managed very easily, also various computation can be performed on the data extracted.

This project uses different Digital Image Processing techniques like Hough Transform, Local Thresholding, Optical Character Recognition, MSER , Canny Edge Detection, morphological operations etc.

We’ve tried to make this solution as general as possible, this approach should work with all kinds of bills, whether it is computer generated bill pictured in low amount of light or handwritten bill image which is not aligned properly.

This project can help with converting piles of bills into soft copies and no further storing is required manually. Also, Receipts/Bills can be stored on cloud and shared. It can also be used in collecting data for surveys.