Named-Entity Recognition in Business Card Images

Shaikh Mohammed Ismail

Department of Computer Science and Engineering
National Institute of Technology Rourkela
Rourkela-769 008, Odisha, India

Named-Entity Recognition in Business Card Images

Thesis submitted in

May 2013

to the department of

Computer Science and Engineering

of

National Institute of Technology Rourkela

in partial fulfillment of the requirements

for the degree of

Bachelor of Technology

in

Computer Science and Engineering

by

Shaikh Mohammed Ismail

[Roll: 109CS0443]

under the guidance of

Prof. Banshidhar Majhi

Department of Computer Science and Engineering
National Institute of Technology Rourkela
Rourkela-769 008, Odisha, India

Computer Science and Engineering
National Institute of Technology Rourkela
Rourkela-769 008, India. `www.nitrkl.ac.in`

Dr. Banshidhar Majhi
Professor

May 09, 2013

Certificate

This is to certify that the work in the project entitled *Named-Entity Recognition in Business Card Images* by *Shaikh Mohammed Ismail* is a record of his work carried out under my supervision and guidance in partial fulfillment of the requirements for the award of the degree of *Bachelor of Technology* in *Computer Science and Engineering.*

Banshidhar Majhi

Acknowledgment

I take this opportunity to express my gratitude and regards to my guide Prof. Banshidhar Majhi for his exemplary guidance, monitoring and constant encouragement throughout the course of this project.

I also take this opportunity to express a deep sense of gratitude to my friends for their support and motivation which helped me in completing this task through its various stages.

I am obliged to the faculty members of the Department of Computer Science & Engineering at NIT Rourkela for the valuable information provided by them in their respective fields. I am grateful for their cooperation during the period of my assignment. In particular I am thankful to Prof. P. K. Sa for offering advice and help on important issues relating to the project.

Lastly, I thank my parents for their constant encouragement without which this assignment would not have been possible.

Shaikh Mohammed Ismail

Abstract

Analysis of document images for information extraction has become very prominent in recent past. Wide variety of information, which has been conventionally stored on paper, is now being converted into electronic form for better storage and intelligent processing. This needs processing of documents using image analysis, processing methods. The objective of Document Image analysis is to recognize the text and graphics components in image of documents and to extract intended information from them.

We are surrounded by text everywhere: window signs, commercial logos and phone numbers plastered on trucks, flyers, take-away menus - and yet to capture and use all this information we essentially resort to typing these phone numbers and websites manually into a phone or computing device. We thought we should help change that, with the help of the mobile phone camera and OCR applications extracting the relevant textual information in these images.

Basically, the problem can be seen as a two step process:

- Extract characters/words from the image by OCR

- Classify the words as Name, Email, Phone No, etc.

Our work was more focussed on the first step - to reduce/minimize the time needed to perform the step given we want to make it usable for mobile computing devices. However, computing under handheld devices involves a number of challenges. Because of the non-contact nature of digital cameras attached to handheld devices, acquired images very often suffer from skew and perspective distortion. In addition to that, manual involvement in the capturing process, uneven and insufficient illumination, and unavailability of sophisticated focusing system yield poor quality images. Since we have to separate text from graphics/background, segmentation/binarization algorithms play a vital role in the process, we studied, analyzed and impelemented existing standard algorithms. A number of thresholding techniques have been previ-

ously proposed using global and local techniques. Global methods apply one threshold to the entire image while local thresholding methods apply different threshold values to different regions of the image. The value is determined by the neighborhood of the pixel to which the thresholding is being applied. An early histogram-based global segmentation algorithm, Otsu's method is widely used. Other techniques use various iterative methods to arrive at a suitable threshold.OCR is done using Tesseract, which is an open-source OCR engine that was developed at HP between 1984 and 1994. Processing follows a traditional step-by-step pipeline.

The second step involves applying appropriate heuristics in order to achieve correct classification. Given a line of text, Named-Entity Recognition(NER) is in itself a different domain of research. We have come up heuristics to identify named-entities, the output of Step 1 is given as input and it displays the information in the relevant field.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

One picture is worth more than ten thousand words. In modern sciences and technologies, images have gained much broader scopes due to the ever growing importance of scientific visualization. With fast computers and signal processors available, digital image processing has become the most common form of image processing and generally, is used because it is not only the most versatile method, but also the cheapest. There is a growing demand of image processing in diverse applications areas such as multimedia computing, secured image data communication, biomedical imaging, biometrics, remote sensing, texture understanding, pattern recognition, content-based image retrieval, compression, and so on.

Analysis of document images for information extraction has become very prominent in recent past. Wide variety of information, which has been conventionally stored on paper, is now being converted into electronic form for better storage and intelligent processing. This needs processing of documents using image analysis, processing methods. Various methods used for digital image processing have three main components: Pre-processing , Feature extraction and the Classification. Pre-processing includes Image acquisition, Binarization, identification, Layout analysis, feature extraction and classification. Classification is an important step in Office Automation, Digital Libraries, and other document image analysis applications.

Document image analysis refers to algorithms and techniques that are applied to

images of documents to obtain a computer-readable description from pixel data. A well-known document image analysis product is the Optical Character Recognition (OCR) software that recognizes characters in a scanned document.

The objective of Document Image analysis is to recognize the text and graphics components in image of documents and to extract intended information from them. Two categories of document image analysis can be defined.

Text Processing deals with the textual components of a document image and its task are:

- Determining the skew (any tilt at which the document may have been scanned in the computer).

- Finding columns, paragraphs, textual lines, words, recognizing the text (possibly its attributes such as size, font etc) by OCR

Graphical processing deals with the non-textual elements (tables, lines, images, symbols, delimiters, company logo etc) Pictures are also included in this category; they are different from graphics in that they are often photographically or artistically generated.

Basic Steps in Document Image Processing

Pre-Processing stage that enhances the quality of the input image & locate the data of interest. It includes

1. Image Acquisition

2. Binarization

3. Noise Reduction

4. Skew Detection and correction

Feature extraction stage that captures the distinctive characteristics of the document under processing.

Classification stage that identifies the document; groups them according to certain classes & helps in their efficient recognition

1.1   Problem Overview

The huge amount of information in all areas of science is overwhelming. If you live active business and social life, contact partners, conduct negotiations, make business deals, you usually exchange business cards. As these cards essentially reflects who the person is with very useful information like Name, email, phone number, website, it seeds a scope of any future communication, if needed. Because of the magnitude of the cards one exchanges each day, it becomes really difficult to store, search and remember important contacts. To ease out the process of remembering and retrieving for future reference, the representation and organization of the information items should provide the user with easy access to the information in which he is interested.

Document image understanding techniques have been widely used in many application domains. Various kinds of documents have been researched and different methods are developed for information retrieval purpose.

Piles of receipts, stacks of business cards, reams of paper, Of course, what we really need is the information trapped inside them. The thesis work is focussed in information extraction from business card images. Given a business card image, our objective is to help people change from manually typing these email address, phone numbers, etc. to an automated one with the help of the mobile phone camera and OCR applications extracting the relevant textual information in these images.

1.2   Applications and Importance

Automatic named-entity recognition has got a number of applications. Starting from having ones information organized such that retrieval can be easy and fast, one can also send the captured card or created contact to any friend.

Adding to the Address Book

The first thing you can do as soon as you get the information from text results is to add the contact to address book, with reviewing and possibly editing the text results.

Send Contact to Friend

Once the contact is created you can send it to your friends so they can have the information with them to help them network with better people.

Network on LinkedIn

Everyone wants to save time and the process to be easy. Using the information in the created contact, you can fire up an invitation to the new contact to join his LinkedIn network.

Web-Sync

The most appealing advantage is that one can sync contact book to a web based account, which might be further synced to multiple devices. Even if you lose your mobile phone, you just need to sign in to your account and download to the new mobile.

1.3   Overview of the Project

Looking closely at the problem it can be considered as a two-step process

- Extract characters/words from the image by OCR

- Classify the words as Name, Email, Phone No, etc.

Although, such applications are commercially available in some recent mobile phones, the accuracy is yet to be extended to be really useful in practice.

Our work was more focussed on the first step - to reduce/minimize the time needed to perform the step given we want to make it usable for mobile computing devices. However, computing under handheld devices involves a number of challenges. Because of the non-contact nature of digital cameras of handheld devices, the images acquired very often suffer from skew and perspective distortion. In addition to that, uneven and insufficient illumination, manual involvement in the capturing process,

and unavailability of sophisticated focusing system yield poor quality images. Since we have to separate text from graphics/background, segmentation/binarization algorithms play a vital role in the process, we studied, analyzed and implemented existing standard algorithms.

1.4   Thesis Organization

The rest of the thesis is organized as follows.

Chapter 2 gives a brief review of basic approaches towards the problem namely thresholding and edge-detection methods. Brief descriptions of existing algorithms and existing methodology for our target system are given.

Chapter 3 describes the proposed Foreground-Background technique.

Chapter 4 presents the output of the algorithms to show the contrast between them in terms of recognition.

Finally **Chapter 5** presents the concluding remarks, with scope for further research work.

Chapter 2

Literature Review

Image segmentation is the process of partitioning a digital image into multiple segments (sets of pixels). The goal of segmentation is to simplify and/or change the representation of an image into something that is more meaningful and easier to analyze. More precisely, image segmentation is the process of assigning a label to every pixel in an image such that pixels with the same label share certain visual characteristics. We focus on two basic forms of Segmentation -

Thresholding methods The simplest method of image segmentation is called the thresholding method. This method is based on a threshold value to turn a grayscale image into a binary image. The key of this method is to select the threshold value (or values when multiple-levels are selected). Image binarization involves separation of pixel values into two groups, black as foreground and white as background. Thresholding plays a major role in binarization of images.

A number of thresholding techniques have been previously proposed using global and local techniques. In images with uniform contrast distribution of background and foreground like document images, global thresholding is more appropriate. In degraded document images, where considerable background noise or variation in contrast and illumination exists, there exists many pixels that cannot be easily classified as foreground or background. In such cases, binarization with local thresholding is more appropriate. **Global methods**(Basic Global Thresholding, Otsu's method) apply one threshold to the entire image while **local thresholding methods** (Niblack, Sauvola, etc.) apply different

threshold values to different regions of the image based on local conditions. The value is determined by the neighborhood of the pixel to which the thresholding is being applied.

Edge detection methods Edge detection is a well-developed field on its own within image processing. Region boundaries and edges are closely related, since there is often a sharp adjustment in intensity at the region boundaries. Edge detection techniques have therefore been used as the base of another segmentation technique. The edges identified by edge detection are often disconnected. To segment an object from an image however, one needs closed region boundaries. The desired edges are the boundaries between such objects. Segmentation methods can also be applied to edges obtained from edge detectors. Lindeberg and Li developed an integrated method that segments edges into straight and curved edge segments for parts-based object recognition, based on a minimum description length (MDL) criterion that was optimized by a split-and-merge-like method with candidate breakpoints obtained from complementary junction cues to obtain more likely points at which to consider partitions into different segments.

Basu et al.[2] presented a novel text/graphics separation methodology for business card images acquired with a cell-phone camera. At first, based on intensity variance, they eliminate the background at a coarse level which made the foreground components distinct from each other. Then using various characteristic features of text and graphics, they remove the non-text components. Finally, the text regions are skew corrected and binarized for further processing.

A number of research works on mobile OCR systems have been found. Mollah et al.[1] present a complete Optical Character Recognition (OCR) system for camera captured image/graphics embedded textual documents for handheld devices. At first, they extract the text regions and skew correct it. Then, they binarize and segment these regions into lines and characters. And then they pass it to character recognition module. Moreover, their developed technique is computationally efficient and consumes low memory so as to be applicable on handheld device.

Shiva et al. [3] proposed a hybrid method to overcome the problem of isolating the foreground text in document images with complex background. According to their three stage proposed approach, In the first stage they identify the candidate text regions using edge detection followed by a connected component analysis. There are chances that because of background complexity the non text regions may get detected as text region. In the second stage they eliminate the false text regions by extracting the texture feature and thereby analyzing the feature value of candidate text regions. In the final stage they separate the text from the background by locally thresholding the image segments narrow down to contain only text. In a nutshell, their approach combines connected component analysis and an unsupervised thresholding to separate foreground text from the complex background in color document images.

Bhaskar et al.[4] presented an algorithm for accurate recognition of text on a business card, given an Android mobile phone camera and image of the card in varying environmental conditions. Reliably interpreting text from real-world photos is a challenging problem due to variations in environmental factors. Even the best open source OCR engine available (Tesseract) gets easily thwarted by uneven illumination, off-angle rotation and perspective, and misaligned text, among others. They showed that by correcting for these factors it is possible to dramatically improve an OCR technologys accuracy. A simple adaptive thresholding technique was used to reduce the effect of illumination gradients. Rotation and perspective were determined by processing a Hough transform. Text alignment was improved by performing vertical segmentation.

First a MATLAB implementation of the algorithm is described where the main objective is to optimize the image for input to the Tesseract OCR (optical character recognition) engine. Then a simplified reduced-complexity implementation on the DROID mobile phone is discussed. The MATLAB implementation is successful in a variety of adverse environmental conditions including variable illumination across the card, varied background surrounding the card, rotation, perspective, and variable horizontal text flow. Direct implementation of the MATLAB algorithm on DROID proved time-intensive, therefore, a simplified version was implemented on the phone.

In summary the paper by Chetan et al. [5] reveals that only few work have been reported towards Graphic separation and Skew correction for Mobile captured documents. In their work they have presented vertical and horizontal projection methods to remove graphics from documents and finally skew correction is performed using Hough transform. Experiments have been performed for documents containing graphics and without graphics and also for noisy images. Experimental results reveal that it works better for both types with graphics and without graphics and therefore proposed method is efficient, novel and accurate for mobile captured documents. Whereas in the case of noisy images the proposed method performance degrades as the noise density increases. A comparative analysis is performed for the existing and proposed method. However proposed method as well as the existing method failed to work for document of low resolution and illumination images.

Garain et al.[6] presented an efficient approach for foreground/background separation in document images with an emphasis on processing of low quality color documents. Algorithm was tested on variety of documents starting from grayscale to color, printed and handwritten manuscripts, documents with well contrasted text as well as those suffering from degradations like uneven illumination, etc. which quite often observed in historical documents. Test documents contained several samples scanned from handwritten manuscripts of famous writers. The manuscripts were written with quill, pencil, etc. and generate low contrast between background and foreground. Results show enormous adaptability of the proposed approach with the uneven illumination or local changes in background and foreground color. The method finds its application in the area of binarization, locating text in documents, image enhancement in digital preservation of ancient documents, compression of documents where foreground and background layers are separated to achieve better compression etc. However, the algorithm has so far been tested on text dominant documents only. Moreover, proper assessment of the extraction results needs benchmarking of foreground and background pixels in sample documents.

2.1 Existing Algorithms

2.1.1 Otsu's Method

In Otsu's method we exhaustively search for the threshold that minimizes the intra-class variance (the variance within the class), defined as a weighted sum of variances of the two classes:

$$\sigma_w^2(t) = w_1(t)\sigma_1^2(t) + w_2(t)\sigma_2^2(t) \tag{2.1}$$

weights w_i are the probabilities of the two classes separated by a threshold t and σ_i^2 variances of these classes.

Otsu shows that minimizing the intra-class variance is the same as maximizing inter-class variance:

$$\sigma_b^2(t) = \sigma^2 - \sigma_w^2(t) = w_1(t)w_2(t)[\mu_1(t) - \mu_2(t)]^2 \tag{2.2}$$

which is expressed in terms of class probabilities w_i and class means μ_i.

The class probability $w_i(t)$ is computed from the histogram as t:

$$w_i(t) = \sum_0^t p(i) \tag{2.3}$$

while the class mean $\mu_1(t)$ is:

$$\mu_1(t) = \left[\sum_0^t p(i)x(i) \right] \Big/ w_1 \tag{2.4}$$

where x(i) is the value at the center of the ith histogram bin. Similarly, you can compute $w_2(t)$ and μ_t on the right-hand side of the histogram for bins greater than t. The class probabilities and class means can be computed iteratively.

2.1.2   Sobel Operator

The operator uses two 3×3 kernels which are convolved with the original image to calculate approximations of the derivatives - one for horizontal changes, and one for vertical. If we define $\mathbf{A}$ as the source image, and G_x and G_y are two images which at each point contain the horizontal and vertical derivative approximations, the computations are as follows:

$$G_x = \begin{bmatrix} +1 & 0 & -1 \\ +2 & 0 & -2 \\ +1 & 0 & -1 \end{bmatrix} * A \tag{2.5}$$

$$G_y = \begin{bmatrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} * A \tag{2.6}$$

where $*$ denotes two 2-dimensional convolution operation.

$$G = \sqrt{G_x^2 + G_y^2} \tag{2.7}$$

2.1.3   Niblack's Method

The method can be described like:

$$pixel = (pixel > mean + k * standard\_deviation)?object : background \tag{2.8}$$

$$T(i,j) = m(i,j) + k * s(i,j) \tag{2.9}$$

Where 'm' is the mean of the number of pixels in that window, 'k' is any constant that can be different for different type of documents and 's' is the standard deviation.

2.1.4 Sauvola's Technique

This efficiently removes the effect of stains in a thresholded image. In our experiments, we used R=128 with 8-bit gray level images and k=0.5 to obtain good results. Below presents the textual binarization formula.

$$T(x,y) = m(x,y).\left[1 + k.\left(\frac{s(x,y)}{R} - 1\right)\right] \tag{2.10}$$

where m(x, y) and s(x, y) are as in Niblack's formula. R is the dynamic range of standard deviation, and the parameter k gets positive values.

Chapter 3

Proposed Work

3.1 Existing Methodology for our Target System[1]

3.1.1 Text Extraction

The input image I is, at first, partitioned into m number of disjoint blocks $B_i, i = 1, 2, ..., m$ such that $B_i \cap B_j = \emptyset$ and $I = \cup_{i=1}^{m} B_i$. Each individual B_i is classified as either Information Block (IB) or Background Block (BB) based on the intensity variation within it. After removal of BBs, adjacent/contiguous IBs constitute isolated components called as regions. These regions are then classified as TR(Text Region) or NR(Non-text region or Graphics region) using various characteristics features of textual and non-textual regions such as dimensions, aspect ratio, information pixel density, region area, coverage ratio, histogram, etc.

3.1.2 Skew Correction

Camera captured images very often suffer from skew and perspective distortion. These occur due to unparallel axes and/or planes at the time of capturing the image. The acquired image does not become uniformly skewed mainly due to perspective distortion. Skewness of different portions of the image may vary between $+\alpha$ to $-\beta$ degrees where both α and β are +ve numbers. Therefore, the image cannot be de-skewed at a single pass. On the other hand, the effect of perspective distortion is distributed throughout the image. Its effect is hardly visible within a small region (e.g. the area of a

Image

Text
Extraction

Skew
Correction

Binarization

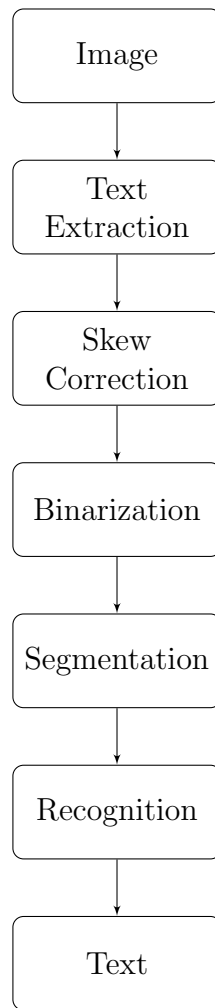Segmentation

Recognition

Text

Figure 3.1: Block Diagram of the present system

character) of the image. So, these text regions are de-skewed using a computationally efficient and fast skew correction technique

3.1.3   Binarization

A skew corrected text region is binarized using a simple yet efficient binarization technique before segmenting it. The algorithms have been described later. In these algorithms, the immediate neighbors around the pixel subject to binarization are also taken as deciding factors for binarization. This type of approach is especially useful to connect the disconnected foreground pixels of a character.

3.1.4   Text Region Segmentation

After binarizing a text region, the horizontal histogram profile $f_i, i = 1, 2, \cdots, H_R$ of the region is analyzed for segmenting the region into text lines. Here f_i denotes the number of black pixel along the ith row of the TR and H_R denotes the height of the de-skewed TR. At first, all possible line segments are determined by thresholding the profile values. The threshold is chosen so as to allow over-segmentation. Text line boundaries are referred by the values of i for which the value of f_i is less than the threshold. Thus, n such segments represent n-1 text lines. After that the inter-segment distances are analyzed and some segments are rejected based on the idea that the distance between two lines in terms of pixels will not be too small and the inter-segment distances are likely to become equal. A detail description of the method is given in [18]. Using vertical histogram profile of each individual text lines, words and characters are segmented.

3.1.5   OCR - Tesseract Algorithm

The Tesseract OCR algorithm has the following steps:

Input A Grayscale or color image is provided as input. The input data should ideally be a *"flat"* image from a flatbed scanner or a near parallel image capture. No rectification capability it provided to correct for perspective distortions

Adaptive thresholding (Otsu's method) performs the reduction of a grayscale image to a binary image. The algorithm assumes that in an image there are foreground (black) pixels and background (white) pixels. It then calculates the optimal threshold that separates the two pixel classes such that the variance between the two is minimal.

Connected-component labeling - Tesseract searches through the image, identifies the foreground pixels, and marks them as *"blobs"* or potential characters.

Line finding algorithm - lines of text are found by analyzing the image space adjacent to potential characters. This algorithm does a Y projection of the binary

image and finds locations having a pixel count less than a specific threshold. These areas are potential lines, and are further analyzed to confirm.

Baseline fitting algorithm - finds baselines for each of the lines. After each line of text is found, Tesseract examines the lines of text to find approximate text height across the line. This process is the first step in determining how to recognize characters.

Fixed pitch detection - the other half of setting up character detection is finding the approximate character width. This allows for the correct incremental extraction of characters as Tesseract walks down a line.

Non-fixed pitch spacing delimiting - characters that are not of uniform width or of a width that agrees with the surrounding neighborhood are reclassified to be processed in an alternate manner.

Word recognition - after finding all of the possible char- acter *"blobs"* in the document, Tesseract does word recognition word by word, on a line by line basis. Words are then passed through a contextual and syntactical analyzer which ensures accurate recognition.

3.2 Proposed Technique

The proposed Foreground-Background technique works as follows:

$$pixel = (pixel > variance - k)?object : background \tag{3.1}$$

$$T(i,j) = var(i,j) - k \tag{3.2}$$

Where 'var' is the variance of the number of pixels in that window, 'k' is a constant that can be different for different type of documents/cards. **In our experimentation we found that business cards responded well for k=100**.

16

Chapter 4

Results and Analysis

Figure 4.1: Output - Otsu's method

Experimenting with different kind of business cards and varying the value of 'k' we found that for k 100, the cards respond well and give better recognition of characters. At the moment our algorithm doesn't give 100% accurate results but it performs better that some of the existing algorithms like - Otsu's , Niblack's and Edge operator but in general Sauvola's technique outperforms our technique.

17

```
Kallol Das

Area Sales Manager

(Higher Academic)

94333 33608 (M)

e~mai|  '  ka||o1.das@oriemblackswancom



Oricm Bluckswan Private Limited
_  '  17,  Chittaranjan  Avenue,  Kulkata-700072
0'19"'  3'ai'k§"'a"  Phones : 2212-8052 / 8054, 3023 4355
Fax :O33-2212-9673

Email :kolkata@orientbLackswan.com
Web :www.or1'entblackswan.c0m
```

Figure 4.2: OCR Result - Otsu's method

Figure 4.3: Output - Sobel Operator

```
K@UH@1] @@@

Ar/@a Saflgs Man@g@v

(IH|ag|?n@r A©@@@ms1@)

@5333 3% (M)

@=maflfl : kauu©II.@as@@vi@mb1Ila@k$w@@.@©m




Qrmiem Bllauekmvamz Pfr'fivZmL@ Lfimfiiced
a '17" Chixttarrannjan Avcemua K@uka(za=7@@@72
©U'D@@3 E"@@[%3§W?MJ Ph®mw@ss : 22n2=©5z2 / (954, M23 @355)

Fax : @331=221:2=@@73

Emafiu :fl<@flDza@a@@fl@mbfla@kssw@w.<:©m
Wceb :www«©n'@W\Ebfl@@|k§wan.E@m
```

Figure 4.4: OCR Result - Sobel Operator

Figure 4.5: Output - Sauvola's method

```
Kallol Das

Area Safles Manager

(Higher Academic)

94333 33608 (M)

email 2 kaIIo1.das@or1entbIackswan.com




Orient Bluckswan Privulc Limited
17, Chittaranjan Avenue, Kolkata-700072
:2212»8052 / 8054, 3023 /+356
2033-2212-9673

: kolkata@orientbLackswan.com
:www.on'entb[ackswan.£am
```

Figure 4.6: OCR Result - Sauvola's Method

Figure 4.7: Output - ForegroundBackground method

```
Kallol Das

Area Sales Manager

(Higher Academic)

94333 33608 (M)

e~mai| ' ka||o1.das@ovienlblackswancom




Orient Bluckswan Private Limited
_ ' 17, Chittaranjan Avenue, Kulkata-700072
0'10"' Bl'=Wk5""""" Phones ; 2212-8052 / 3054, 3023 4355
Fax 1033-2212-9673

Email :kolkata@orientblackswan.com
Web :www.or1'entblackswan.com
```

Figure 4.8: OCR Result - ForegroundBackground method

Figure 4.9: Output - Otsu's method

NGEE ANN

POLYTECHNIC

Ong-Quek Mui Hoon 3]! f; 35-

SENIOR LECTURER
SCHOOL OF INFOCOMM TECHNOLOGV

DD: (65) 6460 6724
Hp: (65) 9028 0010
Email: omh@np.edu.sg

Figure 4.10: OCR Result - Otsu's method

Figure 4.11: Output - Sobel Operator

```
PÂ©iVY[ECHNDC

 NGEE ANN

Â©|71@=@U@k//V\\Mï¬   [H|Â©Â©m EH53 5% %

SENHOR LECTURER
SCHQOE QF ï¬,N|FÂ©CÂ©NJ\A TECHNQLQGY

DH: (<<Â§5B 64360 @724)
Hp: K65]] 9028 mm
Eï¬ rmï¬ ï¬,: @mJh@mxp.@dLw.Â§/gy
```

Figure 4.12: OCR Result - Sobel Operator

Figure 4.13: Output - Sauvola's method

```
NGEE ANN

Ongâ€"Quek Mui Hoon 3F Â§ 3;

SENIOR LECTURER
SCHOOL OF INFOCOMM TECHNOLOGY

DD- (65) 6460 6724
Hp: [65] 9078 0010
Emailâ€˜ omh\Tâ€œâ€˜np edu.sg
```

Figure 4.14: OCR Result - Sauvola's Method

Figure 4.15: Output - ForegroundBackground method

```
Ong-Quek Mui Hoon 3]! Â§= 3?-
SENIOR LECTURER
SCHOOL OF INFOCOMM TECHNOLOGV

DD-, (65) 6460 6724
Hp: (65) 9028 0010
Email: omh@np.edu.sg
```

Figure 4.16: OCR Result - ForegroundBackground method

Chapter 5

Conclusion

This thesis proposes a new thresholding technique based on local variance of a block of pixels from an image. It follows similar framework as by Mollah et al.[1]

And as mentioned earlier, experimenting with different sample of business cards and varying the value of 'k' we found that for k $\sim$ 100, the cards respond well and give better recognition of characters. At the moment our algorithm doesn't give 100% accurate results but it performs better that some of the existing algorithms like - Otsu's , Niblack's and Sobel operator but in general Sauvola's technique outperforms our technique.

5.1 Future Work

We look forward to improve our thresholding technique so that it works for all kinds of documents/cards and to achieve better character recognition so that it can be used as a tool for automatic named-entity recognition from business cards. Though there are applications available by **ScanBizCards** for iPhone, Android, Windows Phone, we are looking for a faster and effective algorithm so that the whole process can be made fast and easy to use for he user.

Bibliography

[1] AF Mollah, S Basu, M Nasipuri, and DK Basu. Design of an optical character recognition system for camera-based handheld devices. *International Journal of Computer Science Issues(IJCSI)*, 8(1), July 2011.

[2] Ayatullah Faruk Mollah, Subhadip Basu, Mita Nasipuri, and Dipak Kumar Basu. Text/graphics separation for business card images for mobile devices. *IAPR International Workshop on Graphics Recognition*, 2009.

[3] Nirmala Shivananda and P. Nagabhushan. Separation of foreground text from complex background in color document images. 2009.

[4] Nicholas Lavassar Sonia Bhaskar and Scott Green. Implementing optical character recognition on the android operating system for business cards. *EE 368 Digital Image Processing*.

[5] H.K.Chethan and G.Hemantha Kumar. Graphics separation and skew correction for mobile captured documents and comparative analysis with existing methods. *International Journal of Computer Applications (0975 8887)*, 7(3), September 2010.

[6] Utpal Garain, Thierry Paquet, and Laurent Heutte. Comparison of some thresholding algorithms for text/background segmentation in difficult document images. *International Journal of Document Analysis and Recognition (IJDAR)*, 7(3), September 2010.