IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 7, Issue 3, March 2018

Smartphone Based Optical Character Recognition for Indian Language

Abhishek R. Katyare[1], Pradeep R. Todkar[2], Shivam C. Bapat[3], Mayur K. Gunjal[4], Prof. Apurva N. Varade[5]

B.E Student, Department of IT, Sinhgad Academy of Engineering, Pune, Maharashtra, India[1,2,3,4]

Assistant Professor, Department of IT, Sinhgad Academy of Engineering, Pune, Maharashtra, India[5]

Abstract: India is a diverse country with various cultural and traditional differences. There are more than 12 distinguished different languages in the country. Smartphone's have been known as most commonly used electronic device in daily life today. As hardware embedded in Smartphone's can perform many more tasks than traditional phones, the Smartphone's are no longer just a communication device but also considered as a powerful computing device which is able to capture images, record videos, etc. With advancement of technology, it is possible to apply some techniques to perform text detection and translation. Therefore, an application that allows Smartphone's to capture an image and extract the text from it to translate into Indian languages is no longer a dream. In this paper, an Android app is proposed by integrating Tesseract OCR engine. Optical Character Recognition (OCR) is designed in a way that automatically identifies the language of the input document for further processing.

Keywords: Android, OCR, text translator, Image Processing.

1. INTRODUCTION

Often when we travel to different places in India, we do not know the script. As a result of which, we travel around in bizarre bubble, cut off from the rest of the society due to the inability to communicate. We are unable to read the signboards and follow the information it conveys. We are molded and shaped by our surroundings and influenced by our mother tongue, due to which it's difficult to cope up with other languages.

Besides that, it is inconvenient for a travelers to carry on their tasks in a foreign country if they don't understand the language used in that country. They need to carry a pocket dictionary or use online translation service in order to understand the message. It is also meant the same way in another study that users are unable to write the text of what they see. This issue might cause a communication breakdown for mankind from a different language background as they are unable to understand the language even though the pocket dictionary and online translation service provided [3]. If the message is unreachable to mankind with different language background, it might cause important information to be missed out [2].

Signboard information plays an important role in our society. Their format is often concise and the information they give is usually very useful. However, the foreign visitors may not understand the language that the signboard is written in, with the consequent loss of all that important information also they pose problems or even tend to be dangerous when we are not familiar with them. The gaining momentum of portable mobile devices, the growing of their computational power and the inclusion of cheap digital cameras on them makes it possible to change from the classical hand dictionary translation to a new faster, comfortable and affordable way. In this sense, it is expected that a high percentage of the world population will own a mobile phone with an embedded camera, which is all that our system needs.

To utilize information present in the textual images, they need to be processed for extracting and recognizing text. Hence, text extraction and recognition is a field that is gaining a lot of research interest nowadays. There are numerous techniques published for text extraction and recognition from scene. Although there exist a lot of research activities in this field, scene text extraction is still remained as a challenging problem, mainly due to two issues: different variety of text patterns like fonts, colors, sizes, orientations; and presence of background outliers similar to text characters, such as windows, bricks and character-like texture [9].

The Product OCR is to provide an efficient and enhanced software tool for the users to perform sign board translation and document image analysis by reading and recognizing the characters and translate them into Indian Language. Irrespective of all fonts and sizes the product will recognize them search them and process them according to the needs of the Environment.

IJARCCE
ISSN (Online) 2278-1021
ISSN (Print) 2319-5940
International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 7, Issue 3, March 2018

II. OPTICAL CHARACTER RECOGNITION

Optical Character Recognition (OCR) is a system that used to recognize the characters on a text. OCR is needed when the information should be readable both to humans and to a machine and alternative inputs cannot be predefined. In comparison with the other techniques for automatic identification, optical character recognition is unique in that it does not require control of the process that produces the information [4]. OCR system can be seen as a combination of several subsystems that have specific ability. The subsystems that are used in OCR on this research are:

a) Image Acquisition

On this research, we use the mobile device as the OCR system platform. The image acquisition process is conducted by using camera on the mobile device. The recorded image captured by the camera is converted by the device to be a digital image so it can be processed.

b) Preprocessing

Preprocessing has an important role to improve the system accuration value. The preprocessing is used to minimize the noise contained by the image resulted from the acquisition process. Binarization preprocessing is one kind of preprocessing step where the image on grayscale level is converted to binary image contained only 1 or 0 value for every pixel.

c) Segmentation

The segmentation process conducted to separate the characters that will be recognized in the next step because the recognition process only run for one character. There are many method that can be used in segmentation process, such as projection, region growing and contour detection.

d) Feature Extraction

Every character has unique features that differentiate it from other characters. To retrieve the features, it needs specific feature extraction method. In the previous research, the features is calculated from the distance of centroid to the contour of the character. The method is simple and rarely used on OCR, but this method resistant to the rotation and resizing.

e) Classification

The feature acquired from the feature extraction process is then classified. One method that often used to classify the features is Backpropagation Neural Network. This method uses the supervised learning process. When a feature given to the network, the weight is updated to reduce the difference of the output pattern with the expected pattern [5].

III. LITERATURE SURVEY

In Scene Text Extraction Using SWT for Tourist Translator on Android Platform paper Miss Pooja Chavare & Dr.Archana Ghotkar has developed an android application which is useful for tourists visiting other countries it helps tourists to roam around the country. The application captures an image which of signboard which is not in readable format of user, it converts/translates the image into the foreign language which user is comfortable with. With this application tourists can recognize book pages, signboards, banners and hotel menus, etc. Application uses SWT for text extraction from image. Author has also used the Tesseract engine of Google for OCR purpose. Considering the Text Format, image can be classified into 2 types

i) Document image
ii) Scene text image.

In his paper Srivastava et al [6] used stroke width and nearest neighbor constraints for text detection. This methods are robust to text regions with weak edges.

A survey on camera captured scene text detection and extraction towards Gurmukhi Script. paper brings light to major challenges faced while extracting the camera-captured scene texts, and scope and applications of text extraction. It also states some different techniques used for text extracting for captured images from camera. This paper provides background, challenges, scope and applications of the camera based text extraction .According to Amandeep Kaur,Renu Dhir,Gurpreet Singh Lehal the challenges can be classified into two groups: Photometric and Geometric

IJARCCE
ISSN (Online) 2278-1021
ISSN (Print) 2319-5940
International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 7, Issue 3, March 2018

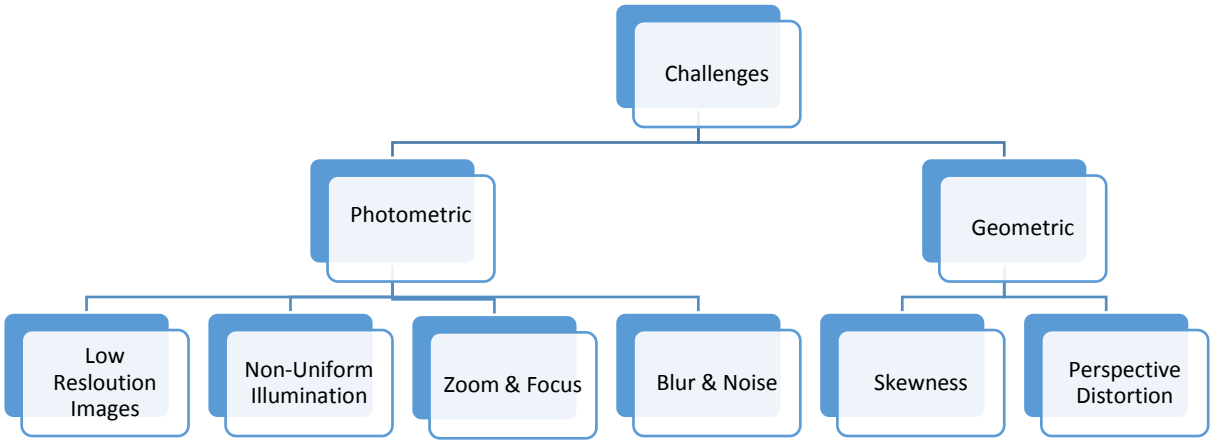


Fig. 3.1. Challenges in Text Extraction

In Recognition Of Devanagari Numericals Using Gabor Filter Shraddha Arya, Indu Chhabra, Gurpreet S. Lehal has used one of the well-known technique for feature extraction called as Gabor Filter. Gabor Filter can capture image characteristics in both time and frequency domain parallel. This paper provide Offline System for feature extraction of Handwritten Devanagari Numericals using Gabor Filter. Author has used Three filter sizes (7x7,19x19,31x31) to find optimal filter size. Firstly the images in the databases are binarized and normalized, after that they are extracted. The Classification is done with the help of Nearest Neighbor and Support Vector Machine. Authors achieved accuracy of 98.06% for recognizing Devanagari numerical [7]. Gabor Filters are widely used in Pattern Recognition, fingerprint recognition, number plate recognition, object tracking, etc.

IV. PROPOSED ARCHITECTURE

The proposed system has to be GUI. It is trained to Indic Script (language), using Tesseract Training Procedures. It will much beneficial for tourists and normal users.
The inputs to the OCR-Engines are:

- Sample Training Images
- Data Files
- Dictionary
- Final trained data

The Image taken from the user goes under following preprocessing steps which helps us to get the image clearer.

A. Gray scaling

Gray scale imaging is sometimes called "black and white," but technically this is a misnomer. In true black and white, also known as halftone, the only possible shades are pure black and pure white. Gray shading in a halftone image is obtained by considering the image as a grid of black dots on a white background (or vice versa), and the sizes of the individual dots will determine the Apparent lightness of the gray in their vicinity. In the case of transmitted light (for example, the image on a computer display), the brightness levels of the red (R), green (G) and blue (B) components are each represented as a number from decimal 0 to 255, or binary 00000000 to 11111111. For every pixel in a red-green-blue (RGB) grayscale image, R = G = B. The lightness of the gray is directly proportional to the number representing the brightness levels of the primary colors. Black is represented by R = G = B = 0 or R = G = B = 00000000, and white is represented by R = G = B = 255 or R = G = B = 11111111. Because there are 8 bits in the binary representation of the gray level, this imaging method is called 8-bit Grayscale. Grayscale can be collectively called as the range of shades of gray. The darkest possible shade is black, which is the total absence of transmitted or reflected light. The lightest possible shade is white, the total transmission or reflection of light at all visible wavelengths. So the very first step in this method is the conversion of the input image i.e. MRI image to be pre-processed into a Grayscale image.

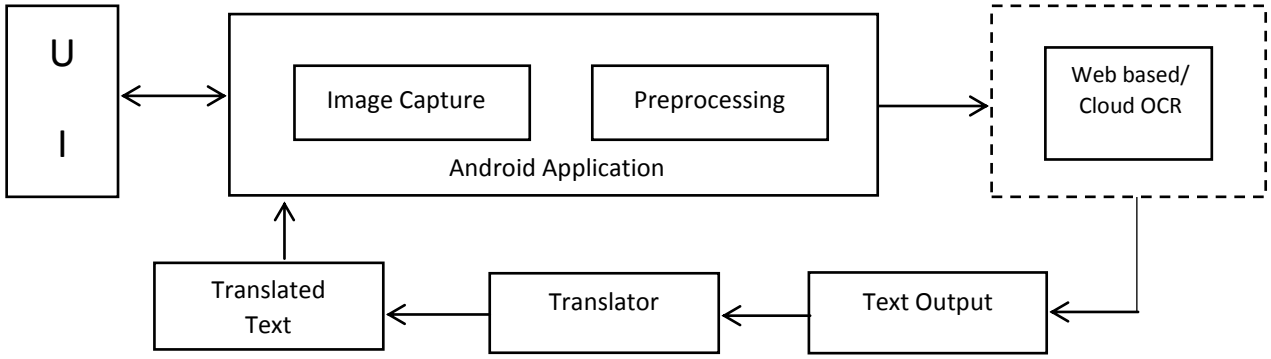
B. Segmentation Thresholding

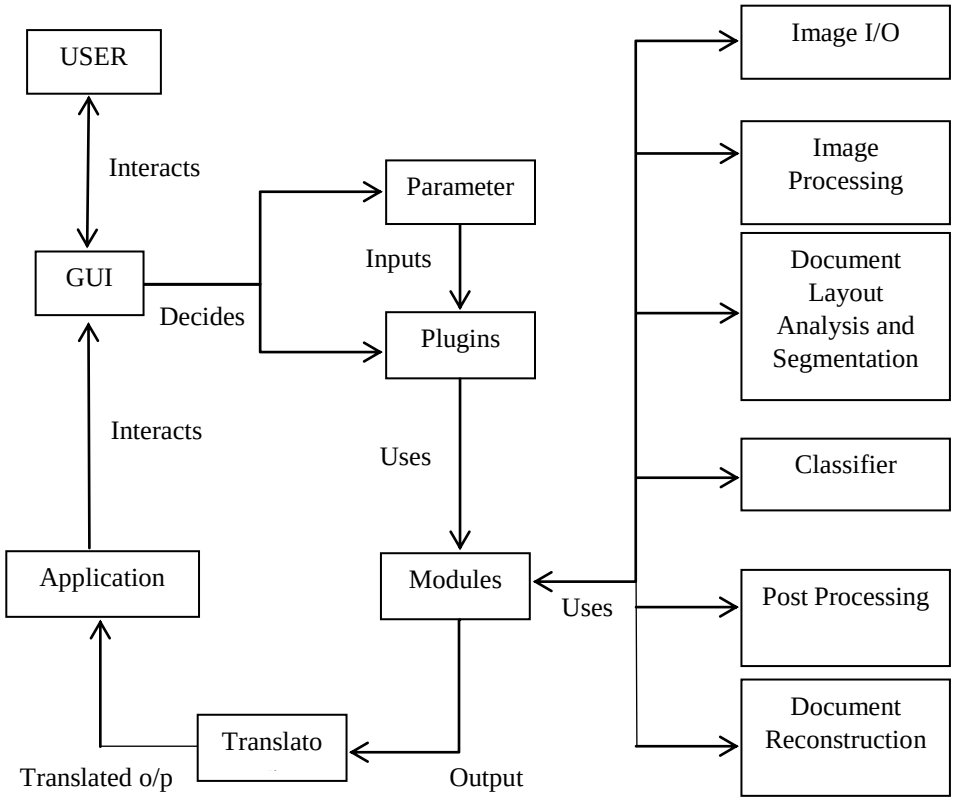
Segmentation is the process of partitioning a digital image into multiple segments (sets of pixels, also known as super pixels). The goal of segmentation is to simplify and/or change the representation of an image into something that is

IJARCCE
ISSN (Online) 2278-1021
ISSN (Print) 2319-5940
International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 7, Issue 3, March 2018

more meaningful and easier to analyze. Image segmentation is typically used to locate objects and boundaries (lines, curves, etc.) in images. 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. Each of the pixels in region are similar with respect to some characteristic or computed property, such as color, intensity, or texture. Adjacent regions are significantly different with respect to the same characteristics. The simplest method of image segmentation is called the thresholding method. This method is based on a threshold value to turn a gray-scale image into a binary image. The key feature of this method is to select the threshold value.

It satisfies all the needs as extensibility, simplicity, interoperability, portability, powerful data structures and also Unicode support. The GUI was developed in android Studio, which is Threaded application. Analysis- proposed system Proposed system is developed as client-server image processing application, having android device as a client and HTTP Apache server. Fig. 4.1 shows proposed system architecture. Using this architecture, user having android mobile will first capture an input image and send the image to a server via HTTP. A PHP script on the server then invokes the server-side application to process the image and perform recognition and translation. After the computation is completed, result will be send back to the android device and display the same on to the screen of the mobile phone.





IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 7, Issue 3, March 2018

V. DISCUSSION

Goal of this project is to get the script form the images, and provide it to user for further processing or translate the extracted script, Extensive research has been done on image classification for different purposes like face recognition, identification of different objects and identification/extraction of text from image having some background. Text identification is an active research area where by system tries to identify the text area in a given image. Text area identified is then passed to OCR system for further recognition of the text. This work is about classifying image area in two classes text and non text using SVM (support vector machine). We identified the features and train a model based on the feature vector which is then used to classify text and non text area in an image. The system reports 70.5% accuracy for caption text images, 70.43% for document text images and 50.40% for scene text image[1].In Scene Text Extraction Using SWT for Tourist Translator on Android Platform ccv's SWT implementation performs on ICDAR 2003 dataset achieved similar performance with what Epshtein et al. reported in their paper, namely, with the old measure method described in ICDAR 2003 contest, ccv's implementation was able to achieve precision rate at 66% and recall rate at 59% (numbers reported in the paper are precision rate 73% and recall rate at 60%)[6].Fingerprint is widely used in identification and verification systems. In Recognition Of Devnagari Numericals Using Gabor Filter, Shraddha Arya,Indu Chhabra and Gurpreet S. Lehal present a novel feature extraction method based on Gabor filter and Recursive Fisher Linear Discriminate (RFLD) algorithm, which is used for fingerprint identification. Experimental results show that applying RFLD to a Gabor filter in four orientations, in comparison with Gabor filter and PCA transform, increases the identification accuracy from 85.2% to 95.2% by nearest cluster center point classifier with Leave-One-Out method. Also, it has shown that applying RFLD to a Gabor filter in four orientations, in comparison with Gabor filter and PCA transform, increases the identification accuracy from 81.9% to 100% by 3NN classifier [7]. The proposed method has lower computational complexity and higher accuracy rates than conventional methods based on texture features.

VI. CONCLUSION AND FUTURE SCOPE

We proposed a system to translate signboard images taken with a mobile phone camera from Indian Languages to English. Since the computational resources of mobile phones are limited, we had to use fast, simple and accurate possible algorithms to work in the most common situations. Our system shows some characteristics that make it interesting and deserve further research: Other systems like Chinese - English translation have been proposed, but no research has been found for Indian Languages to English translation of outside signboard texts.

Proposed android application can be further extended to deal with any target and source language for translation. It can be further modified to deal with text having vertical or arbitrary orientation.

REFERENCES

[1] Amandeep Kaur,Renu Dhir,Gurpreet Singh Lehal, 'A survey on camera captured scene text detection and extraction towards Gurmukhi Script', Springer-Verlag London 2017.
[2] A. Canedo-Rodriguez, S. Kim, J. Kim and Y. Blanco-Fernandez, 'English to Spanish translation of signboard images from mobile phone camera', IEEE Southeast on 2009, 2009.
[3] L. Zhifang, L. Bin and G. Xiaopeng, 'Test automation on mobile device', Proceedings of the 5th Workshop on Automation of Software Test - AST '10, 2010.
[4] A. Joshiy, M. Zhangx, R. Kadmawalay, K. Dantuy, S. Poduriy, and G. S. Sukhatme, "OCRdroid: A Framework to Digitize Text Using Mobile Phones," OCRdroid Project, Retrieved Dec. 10, 2009.
[5] Devinder Singh and Baljit Singh Khehra, "Digit Recognition System Using Back Propagation Neural Network", International Journal of Computer Science and Communication Vol. 2, No. 1, January-June 2011, pp.
[6] Miss Pooja Chavare & Dr.Archana Ghotkar,' Scene Text Extraction Using SWT for Tourist Translator on Android Platform', 2016 International Conference on Automatic Control and Dynamic Optimization Techniques (ICACDOT) International Institute of Information Technology (I²IT), Pune.
[7] Shraddha Arya,Indu Chhabra,Gurpreet S. Lehal, 'Recognition Of Devnagari Numericals Using Gabor Filter', Indian Journal of Science and Technology, Vol 8(27), DOI: 10.17485/ijst/2015/v8i27/81856, October 2015