

Text Extraction from Text Based Image Using Android

Amar G. Waghade¹, Anuja V. Zopate², Ankita G. Titare³, Suraj A. Shelke⁴

¹ Professor, Dept. Of CSE Engineering, DES'sCOET Dhamangaon Railway, Maharashtra, India

^{2,3,4} Student, Dept. Of CSE Engineering, DES'sCOET Dhamangaon Railway, Maharashtra, India

Abstract – In world the smartphones are everywhere and everyone uses smartphones for their day to day life. The smartphones have been developed for various purposes like capturing images, record videos, surf the internet and etc. With advancement of technology, it is possible to apply some techniques to perform text detection and translation. In this study, we have tried to integrate the Tesseract OCR engine and the Google Vision library and develop an application on android platform that allows user to capture the images using camera and extract the text from it. This application recognizes the text that is captured by a mobile phone camera and displays back the recognized text on to the screen. To develop this application we have used the Optical Character Recognition, OCR engine, and we develop our own open source Android application. This application has come up with the solutions for the problems of retyping any hard copy of any documents, scanning the documents in the scanner then use the computer oriented software to recognize the text in the file, or guessing the text and typing it. It aims to develop a piece of software that runs on smartphones and can be used to identify and recognize text of any text based image using the phone's camera. In case of a finding, In order to achieve fast processing times, the project has to deal with the low computing power of smartphones. For this purpose, different approaches to object detection, such as concepts from the machine learning domain, and their implementation on mobile platforms will be analysed.

Key Words: Android, OCR, Text translator, Text to speech.

1. INTRODUCTION

In society, mobile phones over a wide variety of functionalities that are not always related to calling or sending messages. Those functionalities include web browsing, playing games or music, banking, taking photos and so much more. The Text Extraction from Text Based Image is an Android application that aims to allow the user to extract the text from the image and after extracting if user want to translate it into another language he will also translate into other language and also to hear a text contained in a picture that has been taken with a mobile phone. It is an application meant to help those who cannot read a text they encounter, like non-native speakers, the visually impaired and the blind people, estimated at 285 millions in 2010 by the World Health Organization. This project is mainly related to image processing to recognize characters in an image.

Considering text element, image can be classified as: (i) document image and (ii) scene text image. Currently available OCR algorithms have been developed to deal with document images. OCR algorithm expects the input to be black and white image and relatively clean and well structured. In these conditions, the direct application of an OCR engine on a scene text image would result in an almost certain failure. Therefore, the first task while designing application using text information in natural scene image is to transform the image in a way that current OCR engines are able to manage. 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, colours, sizes, orientations and presence of background outliers similar to text characters, such as windows, bricks and character-like texture.

In this project, an android platform based tourist translator application is developed, which is able to extract the text from scene image and translate the same from one language to other. Extraction is performed using stroke width transform (SWT) approach and connected component analysis. Extracted texts are recognized using Google's open source optical character recognition (OCR) engine 'Tesseract' and translated to a target language using Google's translation.

1.1 OCR

OCR technology is allows the conversion of image which is scanned of printed character into text or any other information that user want using android mobile. OCR technology uses three phases first is Scanning of documents as optical images. Next is Recognition which involves converting those images to character streams representing letters of recognized words and the final element used to accessing or storing the text which are already converted. Converted text is nothing but the extracted text. When, the user begins by capturing an image using mobile camera containing text .To convert extracted text into any language text synthesizer is used.

Tesseract is an open source engine developed by HP labs between years 1985 to 1995 and then handed over to Google Inc. in 2006. Tesseract combined with the Leptonica Image Processing library which can read a wide variety of image formats and convert them to text in over 60 languages. It works well on all computer operating system as well as Android and iPhone mobile platform. Due to popularity of

Tesseract being open source engine, there are a lot of academic experiments and OCR software developments conducted successfully. Based on study conducted between OCRAD, GOCR and Tesseract, found out that the Tesseract outperform other open source engines.

2. DESIGN AND IMPLEMENTSTION

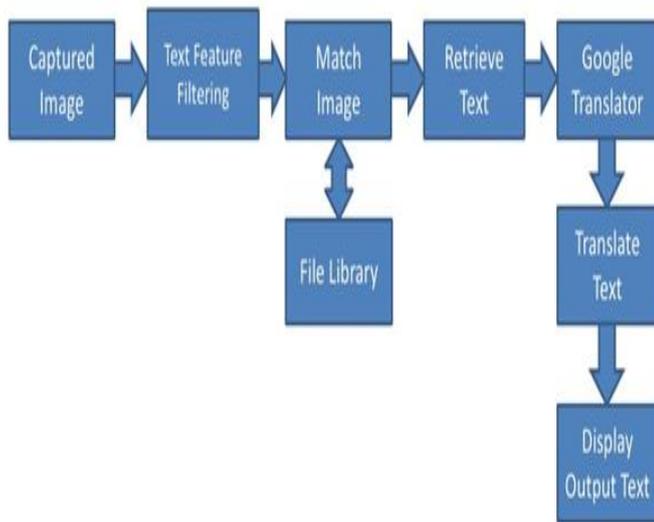


Fig -1:Block Diagram

2.1 Text Feature Filtering

Text feature extraction that extracts text information is an extraction to represent a text message, it is the basis of a large number of text processing. The basic unit of the feature is called text features. Selecting a set of features from some effective ways to reduce the dimension of feature space, the purpose of this process is called feature extraction. During feature extraction, uncorrelated or superfluous features will be deleted. As a method of data preprocessing of learning algorithm, feature extraction can better improve the accuracy of learning algorithm and shorten the time. Selection from the document part can reflect the information on the content words, and the calculation of weight is called the text feature extraction. Common methods of text feature extraction include filtration, fusion, mapping, and clustering method.

2.2 Text-based retrieval

The basic level of multimedia retrieval on the basis of textual information is keyword searching. Numerous methods have been proposed in solving documents indexing and retrieval tasks based on only text content with noisy data, for example the output of an OCR system. These methods can improve the retrieval performance on top of simple word matching using fuzzy logic, confusion information for characters and a bi-gram model, finite state machine, similarity distance measure, and OCR error modeling.

2.3 Google API's

We used the API's in this project like Google translator for translating languages and text to speech engine for converting the text retrieved into the audio. Google APIs is a set of application programming interfaces (APIs) developed by Google which allow communication with Google Services and their integration to other services. Google translator used in this project for translating the present language into any other language that user wants. Google Translate can translate multiple forms of text and media, including text, speech, images, sites, or real-time video, from one language to another.

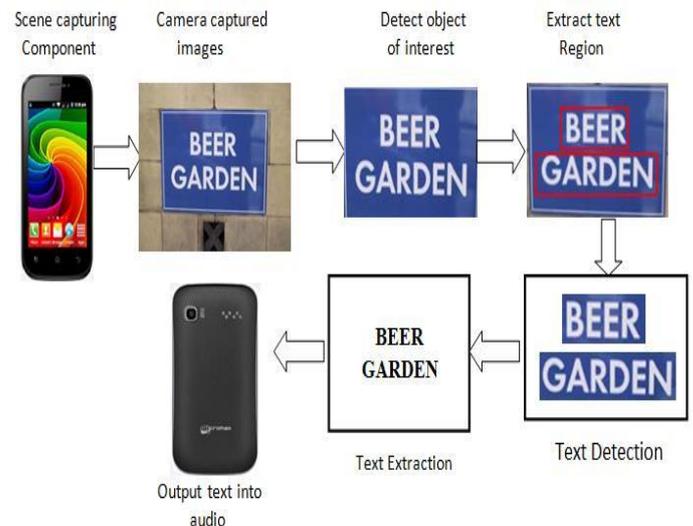


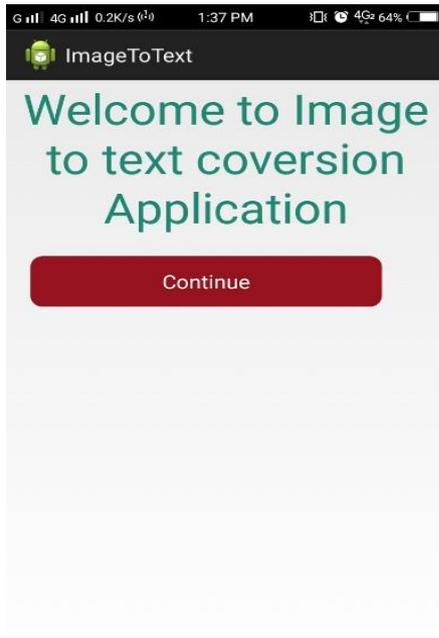
Fig -2: Example of Extracting Text

3. CONCLUSION

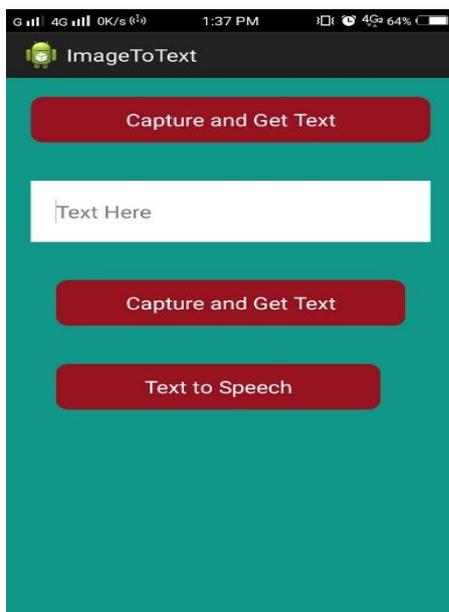
Text extraction is recent research area in the field of Computer Vision. It is challenging problem in the information processing field mainly due to, different variety of text patterns like fonts, colors, sizes, orientations; and presence of background outliers similar to text characters, such as windows, bricks. In this research work, considering few of challenges, real time application named as demo OCR is designed and developed. Extraction is performed using stroke width transform and connected component based approach & using the Tesseract tool using the google vision API. Proposed application assists the tourists, while they are roaming in foreign countries. The performance of system is tested based on extraction rate. With proposed application almost all the text in horizontal orientation extracted correctly, whereas performance of real time images varies with lighting condition and camera resolution.

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.

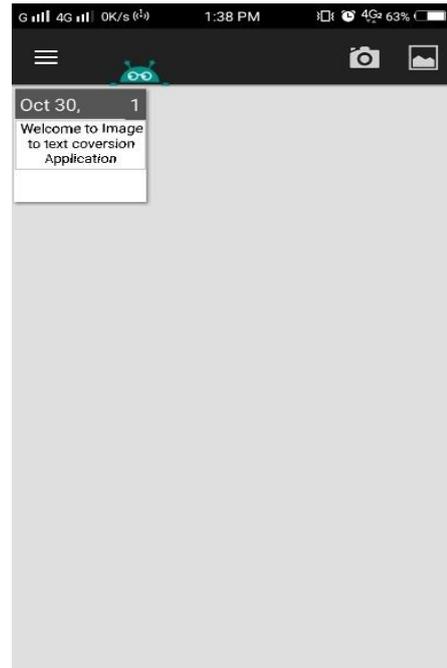
3. SCREENSHOTS



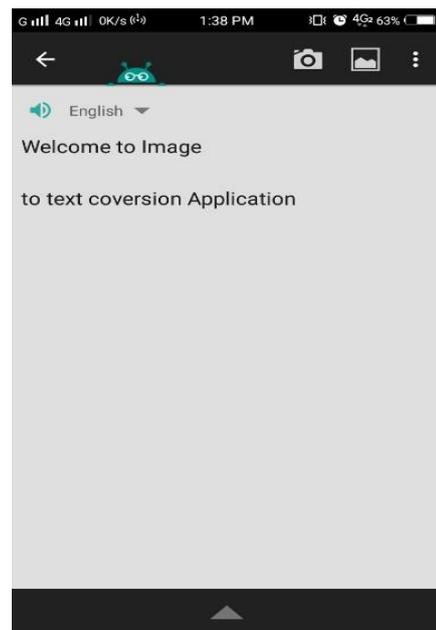
Snapshot 1: StartActivity



Snapshot 2: MainActivity



Snapshot 3: Image to Text



Snapshot 4: Output

REFERENCES

- [1] K. Jung, K. I. Kim, and A. K. Jain, "Text information extraction in images and video: A survey," *Pattern Recognit.*, vol. 37, no. 5, pp. 977-997, 2004.
- [2] X. Zhao, K.-H. Lin, Y. Fu, Y. Hu, Y. Liu, and T. S. Huang, "Text from corners: A novel approach to detect text and caption in videos," *IEEE Trans. Image Process.*, vol. 20, no. 3, pp. 790-799, Mar. 2011.

- [3] W. Kim and C. Kim, "A new approach for overlay text detection and extraction from complex video scene," *IEEE Trans. Image Process.*, vol. 18, no. 2, pp. 401–411, Feb. 2009.
- [4] Jing Zhang and R. Kasturi, "A novel text detection system based on character and link energies" in *IEEE Trans. Image Processing*, Vol. 23, No. 9, Sep 2014.
- [5] Y.-F. Pan, X. Huo, and C.-L. Liu, "A hybrid approach to detect and localize texts in natural scene images," *IEEE Trans. Image Process.*, vol. 20, no. 3, pp. 800–813, Mar. 2010.
- [6] Z. Tu, X. Chen, A. L. Yuille, and S.-C. Zhu, "Image parsing: Unifying segmentation, detection, and recognition," *Int. J. Comput. Vis.*, vol. 63, no. 2, pp. 113– 140, 2005.
- [7] C. Yi and Y. Tian, "Text string detection from natural scenes by structure-based partition and grouping," *IEEE Trans. Image Process.*, vol. 20, no. 9, pp. 2594–2605, Sep. 2011.