

ATMA : Android Travel Mate Application

Ojas Uke, Aqil Shaikh

Department of Computer Engineering, Bharati Vidyapeeth College of Engineering, Mumbai University, Mumbai,
Maharashtra, India

ABSTRACT

Android users are looking for new applications to help them to satisfy their needs. This paper presents an extremely on-demand, speedy Android Application ATMA. ATMA stands for Android Travel Mate Application.. It enables Travellers to easily capture the native country language hand written or printed text. The built-in OCR converts the text embedded in the captured image into Unicode text format and also gives the translation of the text into their desired language. This Application has an advanced search option linked to Google, to search for the whereabouts of the thing. The OCR and Image Processing tools are built in applications. It provides speedy, robust and high Quality performance because of having better Auto focus behaviour, continuous dynamic preview and better noise tolerance feature.

Keywords : Android, Tesseract OCR engine, Bing Translator API, Shirorekha Cutting, Tesseract Android Tools, Android SDK, Android NDK, Tess-two, Eclipse, ADT plug in.

I. INTRODUCTION

The next upcoming operating systems are not on desktops but on the small mobile devices people carry every day. The openness of these new environments leads to new applications & markets and enables greater integration. As the demand grows for mobile phone applications, research in optical character recognition, a technology well developed for scanned documents, is shifting focus to the recognition of text embedded in digital photographs.[1]. Optical character recognition (OCR) is a powerful tool for bringing information from our normal lives into the increasingly digital world. This technology has long seen use in building digital libraries, identifying text from natural scenes, understanding hand-written office forms etc. [1]

By 2014 mobile internet usage should take over Desktop internet usage. Google's approach is to develop an operating system which can run on every mobile device and not for their specific mobile devices itself, enables them to reach as many people as possible.

Every day a Smartphone user may look for a new application dedicated for his need. Android makes it easier for users to get and use new content and applications on their Smart phones. Our project presents an extremely on-demand, fast and user friendly Android Application. This application is useful for native Tourists and Travellers who possess Android Smart phones. This Application has an advanced search feature so that identified as well as translated text can be used to copy, paste, share and search for travel related searches like museums, places, restaurants, books, culture, hotels, etc. There is no remote computing overhead because the application has inbuilt OCR suite as well as Image Processing suite both installed in the Android device. It provides speedy, robust and extremely high Quality performance because of having improved Auto focus behaviour, continuous dynamic preview and better noise tolerance feature. The motivation of a real time text translation mobile application is to help tourists travel in a foreign language environment. The application we developed enables the users to get text translate as ease as a button click. The camera captures the text and returns the translated result in real time.[2]

II. METHODS AND MATERIAL

A. Literature Survey

In Indian market there was always a high demand of such an Android Application, which would enable a Tourist sitting in a restaurant to capture, OCR and translate the Hindi language menu in his own country language to order his favourite dish without any need to know the language of menu. ATMA enables people to understand any particular country language and also seamlessly fire a travel specific search to get the information regarding the country like museums, restaurants, hotels, culture, temples, books, movies, songs, etc. Image processing on mobile phones is a new and exciting field with many challenges due to limited hardware, limited CPU utilization, image Quality and connectivity. Phones with high quality cameras, powerful CPUs, and memory storage devices are becoming increasingly common. Most of the existing OCR technology for smart phones are likely to be remotely executed by the servers which do not face limited memory and limited processing power challenge moreover they suffer from networking delay. The proposed application would have image processing suite as well as OCR engine both installed on the smart phone implying no server communication overhead and quick and far better processing thus would not just only tackle the limited processing power and limited memory challenge but also provide improved performance. [2]

Optical Character Recognition

OCR, Optical Character Recognition, is developed to translate scanned images of handwritten, typewritten or printed text into machine-encoded text. A lot of OCR software has been developed to accomplish this mission. Tesseract, originally developed as proprietary software at Hewlett- Packard between 1985 and 1995, now sponsored by Google, is considered to be one of the most accurate open source OCR engine currently available. It is capable of recognizing text in variety of languages in a binary image format.[3]

Android Technology

The term “Android” has its origin in the Greek word andr-, meaning “man or male” and the suffix -aides, used to mean “alike or of the species”. This together

means as much as “being human”. Android is a software stack for mobile devices which means a reference to a set of system programs or a set of application programs that form a complete system. This software platform provides a foundation for applications just like a real working platform. Android is intended to revolutionize the mobile market by bringing the internet to the cell phone and allowing its use in the same way as on the PC. Android is an operating system based on Linux with a Java programming interface. It provides tools, e.g. a compiler, debugger and a device emulator as well as its own Java Virtual machine (Dalvik Virtual Machine - DVM). Android is created by the Open Handset Alliance which is led by Google. Android uses a special virtual machine, e.g. the Dalvik[2]

B. Existing System

Android: Android is a Linux-based operating system designed primarily for touch screen mobile devices such as smart phones and tablet computers. Initially developed by Android, Inc., which Google backed financially and later bought in 2005, Android was unveiled in 2007 along with the founding of the Open Handset Alliance: a consortium of hardware, software, and telecommunication companies devoted to advancing open standards for mobile devices. The Android Architecture layer is as follows

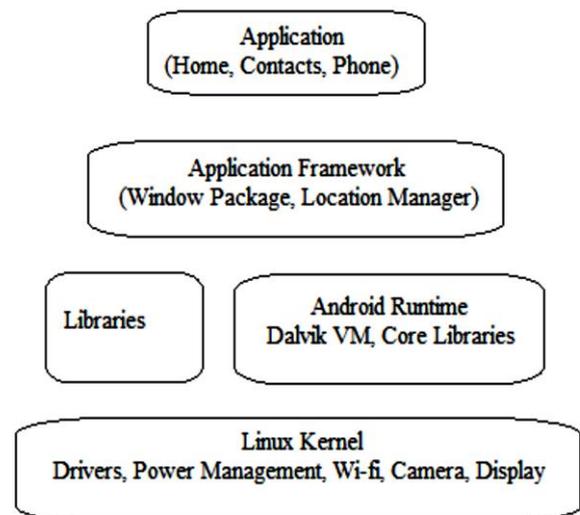


Figure 1. Android Architecture

The Fig 1 shows an open development platform, Android offers developers the ability to build extremely rich and innovative applications. Developers are free to take advantage of the device hardware, access location

information, run background services, set alarms, add notifications to the status bar, and much, much more. Developers have full access to the same framework APIs used by the core applications. The application architecture is designed to simplify the reuse of components; any application can publish its capabilities and any other application may then make use of those capabilities (subject to security constraints enforced by the framework). This same mechanism allows components to be replaced by the user. Underlying all applications is a set of services and systems, including:

Dalvik Virtual Machine

The Dalvik virtual machine as shown in Fig 2, is simple Java interpreter machine, completely optimized for Android platform and which is developed to run on low-end memory mobile devices. One of the prominent aspects in Dalvik is its capability to run along an application compilation enhancing the runtime performance of the applications. DVM delivers a good performance over various stages of an application runtime environment, conserving more battery-power during long run of an application. A small drawback with Dalvik is its non-compatibility with Java SE libraries, Java Applet, Java ME class libraries and Swing Java libraries.[3]

a) Tesseract OCR engine: Tesseract is a free software optical character recognition engine for various operating systems. Tesseract is probably the most accurate open source OCR engine available. Combined with the Leptonica Image Processing Library it can read a wide variety of image formats and convert them to text in over 60 languages. It was one of the top 3 engines in the 1995 UNLV Accuracy test. Between 1995 and 2006 it had little work done on it, but since then it has been improved extensively by Google. It is released under the Apache License 2.0.

b) Tesseract Android Tools: Tesseract Tools for Android provides a set of Android APIs and build files for the Tesseract OCR and Leptonica image processing libraries. Android SDK: The Android SDK provides you the API libraries and developer tools necessary to build, test, and debug apps for Android.

c) ADT: Android Development Tools (ADT) is a plugin for the Eclipse IDE that is designed to give you a

powerful, integrated environment in which to build Android applications.

d) Eclipse: In computer programming, Eclipse is a multi-language software development environment comprising a base workspace and an extensible plug-in system for customizing the environment. It is written mostly in Java. It can be used to develop applications in Java and, by means of various plug-ins, other programming languages including Ada, C, C++, COBOL, FORTRAN, Haskell, Perl, PHP, and Python.

C. Objective

OCR the native country language printed or handwritten image. Translate the Recognized text into one of 4-5 languages. This application enables people to understand any native country language. The proposed application would have image processing suite as well as OCR[6] engine both installed on the smart phone implying no server communication overhead and quick and far better processing thus would not just only tackle the limited processing power and limited memory challenge but also provide improved performance.

D. Proposed System

The project enables travelers to easily capture the native country language hand written or printed text and translate them into unique code text and also provides the translation of the obtained text. It also provides the facility of search engine so that the travelers can obtain the necessary information on the whereabouts, the history and the significance of the event or place. This would prove enormously beneficial with respect to the aspects about localization being a common phenomenon now-a-days. Also android platform has been increasingly being common in accordance with its features like low-cost, customizable, lightweight operating system and more. This application enables people to understand any native country language and also seamlessly fire a travel specific search to get the info regarding the country like museums, restaurants, hotels, culture, temples, books, movies, songs, etc. It overcomes the existing problems with OCR technology i.e. limited memory and limited processing power challenge moreover they suffer from networking delay. In this module, The Binarization of Captured Image takes place, after that the text layout is analyzed, Blobs

are detected and finally words and lines are detected. The words are sent to a number of passes. In these passes each word is chopped into characters and characters are checked for the need of joining the broken characters or the breaking of associated characters. Finally chopped characters are recognized with the help of inbuilt fuzzy features matched to language specific training data of Unicode characters. After each pass the words are matched back and forth with the Language specific Dictionary words.

A. Preprocessing It smooth's the digitized characters to remove noise. The smoothing implies both filling and thinning. Filling eliminates small breaks, gaps and holes in the digitized characters. Thinning reduces the width of the line. The normalization is applied to obtain characters of uniform size, slant and rotation.

B. Segmentation Segmentation is a classifier which helps to fragment each character from a word present in a given image / page. The objective of the segmentation is to extract each character from the text present in the image. After performing Segmentation, the characters of the string will be separated and it will be used for further processing.[4]

C. Feature Extraction In this stage, the features of the characters that are crucial for classifying them at recognition stage are extracted.[7]

D. Classification The classification is the process of identifying each character and assigning to it the correct.

III. RESULTS AND DISCUSSION

A. Implementation

It will be very helpful for travellers to understand & adapt local languages with ease. This Application also provides the facility of search engine so that the travellers can obtain the necessary information on the whereabouts, the history and the significance of the event or place. The application enables the users to get text translate as ease as a button click. The camera captures the text and returns the translated result in real time. This application enables people to understand any native country language and also seamlessly fire a travel specific search to get the info regarding the country like museums, restaurants, hotels, culture, temples, books,

movies, songs, etc. It overcomes the existing problems with OCR technology i.e. limited memory and limited processing power challenge moreover they suffer from networking delay.

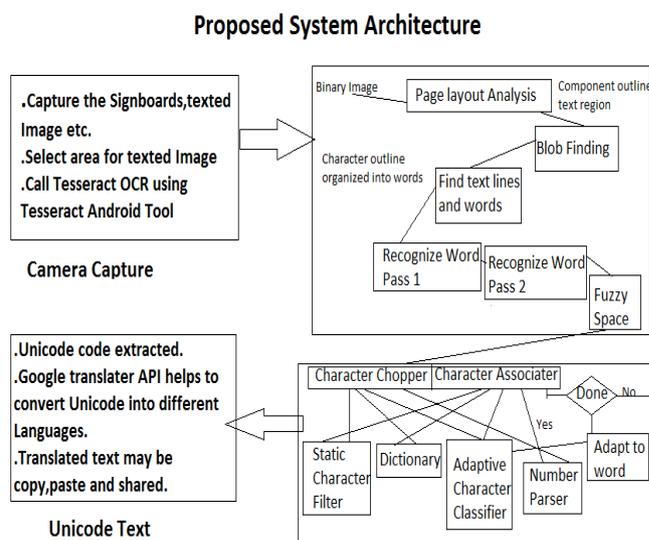


Figure 3. Tesseract Architecture

Fig 3 shows the application consists of four major components described below:

Camera Capture Module

In this module the user is allowed to resize the camera capture box by touching the box corners on the screen so as to capture the only concerned text image from signboard, banner and book pages. The camera keeps continue auto focusing the image automatically throughout the session. Once the capture button is pressed the beep sound plays and the captured image is sent to Tesseract OCR engine module.[4]

Tesseract OCR Engine Module

In this module, The Binarization of Captured Image takes place, after that the text layout is analyzed, Blobs are detected and finally words and lines are detected. The words are sent to a number of passes. In these passes each word is chopped into characters and characters are checked for the need of joining the broken characters or the breaking of associated characters. Finally chopped characters are recognized with the help of inbuilt fuzzy features matched to language specific training data of Unicode characters. After each pass the words are matched back and forth with the Language specific Dictionary words.[7]

B. Results



Figure 1. Main Activity



Figure 2. Home Screen

Fig 1 and 2 shows the main loading/booting screen of the android application as well as the home screen of the app.



Figure 3. File Browser

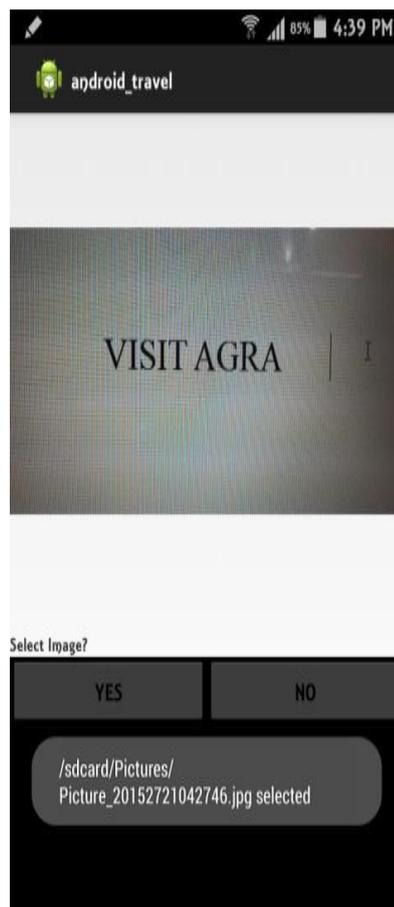


Figure 5. Picture Viewer

Fig 3 represents the file browser history where all the image files have been stored in our SD card/memory of the device in which the application is running.

Fig 4 shows how the image is being captured by the phone/device's camera.



Figure 5. Text to be recognized

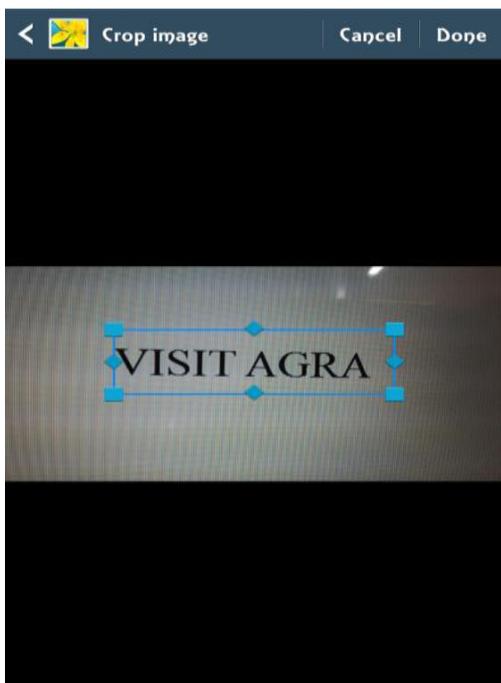


Figure 6. Captured area Selection BoX

Fig 5 and 6 shows how the text has been recognized by the OCR engine and how the cropping function works using the application.



Figure 7. OCR Performed

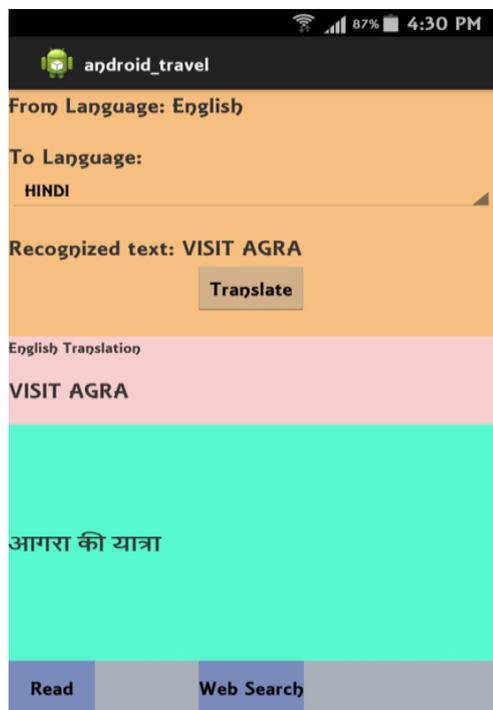


Figure 8. Translation in HINDI

Fig 7 shows how the language translation has been carried out with the help of the OCR engine.

Fig 8 shows how the translation of the character takes place using the OCR engine.

IV. ACKNOWLEDGMENT

We would like to thank our professor Mr. Sandip Chavan of Department of Computer Engineering, BVCOE Kharghar, Navi Mumbai, Maharashtra, India, for providing their support for this project. We would also like to thank our HOD Dr. D.R.Ingle for his unending faith and blessings.

V. CONCLUSION

This paper presented a comprehensive survey on existing OCR Technique by categorizing them according to the features used in each stage. Comparisons of them in terms of pros, cons, recognition results and processing speed were addressed. A future forecast of an OCR is also discussed. We focus on using orientation sensor, embedded high resolution camera and digital image processing techniques to solve OCR related issues on camera captured images. This Application provides fast, robust and extremely high Quality performance because of having improved Auto focus behaviour also provides the facility of search engine so that the travellers can obtain the necessary information on the whereabouts, the history and the significance of the event or place, continuous dynamic preview, improved noise tolerance feature and no remote computing overhead. The future result of OCR should concentrate on multi style plate recognition, video based images using temporal information, 3 D images processing, ambiguous character recognition and so on.

VI. REFERENCES

- [1] Smith, R. "An Overview of the Tesseract OCR" in proc. ICDAR 2007, Curitiba, Paraná, Brazil
- [2] Bansal, V. and Sinha, R.M.K. "A Complete OCR for Printed Hindi Text in Devanagari Script", Sixth International Conference on Document Analysis and Recognition, IEEE Publication, Seattle USA, 2001, Page(s):800-804.
- [3] Saba, T., Sulong, G. and Rehman, A. "A Survey on Methods and Strategies on Touched Characters Segmentation", International Journal of Research and Reviews in Computer Science (IJRRCS) Vol. 1, No. 2, June 2010. 4Pal, U., Chaudhuri, B. B. "Indian Script Character recognition: A survey", Pattern Recognition, vol. 37, pp. 1887-1899, 2004.
- [4] Jindal, M.K., Sharma, R.K., Ingle, G.S. "A Study of Different Kinds of Degradation in Printed Gurmukhi Script", Proceedings of the International Conference on Computing: Theory and Applications (ICCTA'07), 2007.
- [5] Google code : <http://googlecode.blogspot.com/2006/08/announcing-tesseract-ocr.html> (last accessed 8 January, 2012)
- [6] Open source Tesseract OCR Engine, available at <http://code.google.com/p/tesseract-ocr/>
- [7] Tesseract OCR training data downloads. Available at <http://code.google.com/p/tesseract-ocr/download/list>