Android Optical Character Recognition

Shital Malhar¹, Manasi Gosavi², Pooja Lad³
¹,²,³ Dilkap Research Institute of Engineering and Management Studies, Neral.
(Mumbai University)

Abstract - The next generation open operating systems are not on desktops or mainframes but on the small mobile devices people carry every day. The openness of these new environments leads to new applications and markets and enables greater integration. Every day a Smartphone user may look for a new application dedicated for his need. Android makes it easier for consumers to get and use new content and applications on their Smart phones. The Proposed 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. One of the feature in application is it enables Travellers and Tourists to easily capture the native country language Books pages, signboards, banners and hotel menus, currency convertor. The built-in OCR converts the text embedded in the captured image into Unicode text format. It also provides translation facility so that Tourists can translate the Native Language Unicode text into their own country language. There is no remote computing overhead because the application has built in OCR suite as well as Image processing suite both installed in the Android device. The main objective of this application is to help tourist to travel/navigate easily and freely in travelling place without any difficulty.

1. Introduction

The combination of the smart phone and the Internet service is the trend of the future information development and software applications. Mobile phones are the most commonly used communication tools. The motivation of the application is that it enables the users to get English text translated, as ease as a button click. The camera captures the English Text Image and returns the English Text and then one can translate the English text result in 22 languages. Optical character recognition (OCR) has come out with the new open and comprehensive platform for mobile devices called Android. Android includes an operating system, middleware, user-interface and applications. By now mobile internet usage has taken 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. This enables them to reach as many people as possible. This application is useful for native Tourists and Travellers who possess Android Smart phones. The main objective of this application is to help tourist to travel easily and freely without any difficulty. The proposed application also provides currency convertor, a real market value.

1.1 Problem Definition

The Project presents an extremely on-demand, fast and user friendly Android Application. This application is useful for native Tourists and Travelers who possess Android Smart phones. Using mobile phones to obtain information is not only quick, but also more convenient shortcut to improve people's lives. The Proposed Android Travel Mate Application enables Tourists and Travelers so that he/she can:

OCR the native country language Books pages, Signboards, Banners and hotel menus etc. (English)

Translate the Recognized English text into one of 22 National & International languages, which will be helpful in many situations such as in hotels, museums, signboards, notice, warning, instructions, etc. User can see the currency conversion of the destination city currency from its residence country.

1.2 Scope

Applications of OCR include to recognize the English text from images by just clicking the image from camera and to translate that recognized English text to 22 National & International languages, which will be helpful in many situations such as in hotels, museums, signboards, notice, warning, instructions, etc. User can see the currency conversion of the destination city currency from its residence country.
2 Review of Literature

2.1 Android

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. 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). To simplify development Google provides the Android Development Tools (ADT) for Eclipse. Every Android application runs in its own process and under its own user id which is generated automatically by the Android system during deployment. Therefore the application is isolated from other running applications and a misbehaving application cannot easily harm other Android applications.

2.2 Android Architecture

The Android Architecture layers are as follows:

![Android Architecture Diagram](image)

Fig-1: Android Architecture

Android comes with a set of built-in core applications including an email client, SMS program, calendar, maps, browser, contacts, and others which are written using Java. By providing 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).

Android includes a set of C/C++ libraries used by various components of the Android system. These capabilities are exposed to developers through the Android application framework.

2.3 Features of Android

The Android continues to lead the Smart Phone market because of having a great number of features. Android allows one to access core mobile device functionality through standard API available open-source. One can combine information from the web with data on the phone such as contacts or geographic location to create new user experiences. Android does not differentiate between the phones basic and third party applications even the dialler or home screen can be replaced. The SDK contains what one need to build and run Android applications, including a true device emulator and advanced debugging tools. Thus provides fast and easy development. It enables reuse and replacement of components. Dalvik virtual machine is optimized for mobile devices. Integrated browser based on the open source Web Kit engine. Optimized graphics powered by a custom 2D graphics library; 3D graphics based on the OpenGL ES 1.0 specification (hardware acceleration optional). SQLite is used for structured data storage. Media support for common audio, video, and still image formats (MPEG4, H.264, MP3, AAC, AMR, JPEG, PNG, GIF). Android supports GSM Telephony (hardware dependent), Bluetooth, EDGE, 3G, and Wi-Fi (hardware dependent), Camera, GPS, compass, and accelerometer (hardware dependent) etc. Android possesses a rich development environment including a device emulator, tools for debugging, memory and performance profiling, and a plug in for the Eclipse IDE.

2.4 Tesseract OCR Engine

Today, Tesseract is considered one of the most accurate open source OCR engines available. Tesseract OCR Engine was one of the best 3 engines in 1995 UNLV Accuracy Test. Between 1995 and 2006 however; there was little activity in Tesseract, until it was open sourced by HP and UNLV in 2005. It was again re-released to the open source community in August of 2006 by Google [1]. Tesseract has ability to train for newer language and scripts as well [2]. A complete overview of Tesseract OCR engine can be found in [3]. While Tesseract was originally developed for English, it has since been extended to recognize French, Italian, Catalan, Czech, Danish, Polish, Bulgarian, Russian, Greek, Korean, Spanish, Japanese, Dutch, Chinese, Indonesian, Swedish, German, Thai, Arabic, and Hindi etc [4]. Training the Tesseract OCR Engine for Hindi language requires in-depth knowledge of Devanagari script in order to collect the character set.
2.4.1 Components of Tesseract OCR Engine

The application consists of four major components described below:

In camera capture 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 [1]. 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.

In Tesseract OCR Engine 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.

In dictionary words matching module each group of sequential characters is searched for a dictionary based word match, which helps in identifying the word more accurately rather than just giving an meaningless word as result. Finally the recognized text is transferred to Unicode text Post processing Module.

In Unicode text post processing module, the recognized characters are displayed as Unicode characters and the user is allowed to translate the recognized text into his desired language available in the drop down list from settings. From there user may choose one of the two available translators from the drop down list including Bing Translator and Google translator [6]. Moreover the user can use the advanced search feature to search the travel specific related queries like museums, books, videos, songs, culture, images, places and hotels etc. related to recognized or translated text. Open source Tesseract OCR Engine is available on Google [7].

3. Methodology

The proposed Project will be done using android programming and it will be an android application. We will be using client/server architecture to implement the project. There will be 1 server, where admin will be maintaining the user details, user registrations, recommended places to be modified. We will be using php language for it. And there will be 1 client side, where user can access the functionality of application, which will be in android.

3.1 Client and server role

The client–server characteristic describes the relationship of cooperating programs in an application. The server component provides a function or service to one or many clients, which initiate requests for such services.

Servers are classified by the services they provide. For instance, a web server serves web pages and a file server serves computer files. A shared resource may be any of the server computer's software and electronic components, from programs and data to processors and storage devices. The sharing of resources of a server constitutes a service. Whether a computer is a client, a server, or both, is determined by the nature of the application that requires the service functions. For example, a single computer can run web server and file server software at the same time to serve different data to clients making different kinds of requests. Client software can also communicate with server software within the same computer.[8] Communication between servers, such as to synchronize data, is sometimes called inter-server or server-to-server communication.

3.2 Client - server communication

In general, a service is an abstraction of computer resources and a client does not have to be concerned with how the server performs while fulfilling the request and delivering the response. The client only has to understand the response based on the well-known application protocol, i.e. the content and the formatting of the data for the requested service.

4. Conclusion

The friendly tourist application named as “Travel Guru”, an android application as travel mate, has been proposed which includes features such as OCR with translation to 22 languages (Indian as well as International Languages), visited places on map, favourite places, currency convertor, notification specifying the recommended place that user should visit.

5. References


