

# FINAL PROJECT

# OCR TRANSLATOR



## **Release Date:**

04<sup>th</sup> March, 2015

## **Project Advisor**

Mr. Imran Saleem

## **Submitted By:**

Amir Sattar	12025027048
Muhammad Ismail	12025027015
Suneela Zulfiqar	112327030
Shahzad Shafique	12025027005
Muhammad Afzal	12025027007

**School Of Professional Advancement**  
**University of Management and Technology**

---

# OCR Translator

*Provide Best Opportunity to Control*

*Languages Problems*

Project submitted to the Faculty of School of Professional Advancement,  
University of Management and Technology, Lahore Pakistan.

In partial fulfillment of requirement for the degree of

**Master of Computer Science**

**(MCS)**

Approved By

---

**Advisor**

**Mr. Imran Saleem**

---

**Program Head**

---

---

# ABSTRACT

---

The major focus of our application is to translate languages into native language for the ease of tourist and travelers. This app will help tourists to navigate easily in a foreign environment. Most of technologies such as electronic dictionary and other translation services cost traveler's time and effort to retrieve the meaning of certain words. Hence, those travelers are looking for alternative way for translating the text contents, which can be detected by machine translation services for mobile devices. To address this problem, we are developing the mobile phone application that capable of acquiring an image of the text, locating the region of interest (where text is specified) within the image, and translate extracted text into user-desired language.

---

---

# **DEDICATION**

---

To My dearest Parents, My Family, Friends and my Respected Teachers who motivated, supported and encouraged us in every aspect of my life.

---

---

## ACKNOWLEDGEMENTS

---

We would like to thank our project supervisor, **Mr Imran Salem**, his assistance and understanding of our situation that enabled us to complete the report.

We would also give a special word of thanks to **Program Head**, for their help and encouragement over the last three month. Both of them instructed us to carry the work flow in right direction and time to achieve the best result in limited time. We are extremely grateful and remain obliged both of them. We wish to thank our families and friends for their support and prayers over the years. And finally you, whoever you are, for taking the time to read this report.

---

---

# TABLE OF CONTENTS

---

Chapter 1 .....	<b>Error! Bookmark not defined.</b>
1. INTRODUCTION TO THE PROBLEM .....	<b>Error! Bookmark not defined.</b>
1.2.Introduction .....	<b>Error! Bookmark not defined.</b>
1.3.Project Scope .....	<b>Error! Bookmark not defined.</b>
1.4.Objective/Motivation .....	<b>Error! Bookmark not defined.</b>
1.5.Problem Statement .....	<b>Error! Bookmark not defined.</b>
1.6.Description .....	<b>Error! Bookmark not defined.</b>
1.7.Purpose of Documents .....	<b>Error! Bookmark not defined.</b>
Chapter: 2.....	<b>Error! Bookmark not defined.</b>
2.EXISTING SOLUTION AND PROPOSED SOLUTION ...	<b>Error! Bookmark not defined.</b>
2.1 Existing Solution.....	<b>Error! Bookmark not defined.</b>
2.2 Proposed Solution .....	<b>Error! Bookmark not defined.</b>
2.2.1 Scanning.....	<b>Error! Bookmark not defined.</b>
2.2.2 Recognition .....	<b>Error! Bookmark not defined.</b>
2.2.3 Transformed Text.....	<b>Error! Bookmark not defined.</b>
Chapter: 3.....	<b>Error! Bookmark not defined.</b>
3.ANALYSIS DOCUMENTS .....	<b>Error! Bookmark not defined.</b>
3.1.Vision Documents.....	<b>Error! Bookmark not defined.</b>
3.1.1Optical Character Recognition.....	<b>Error! Bookmark not defined.</b>
3.1.2 Function of OCR.....	<b>Error! Bookmark not defined.</b>
3.1.3 Features of OCR .....	<b>Error! Bookmark not defined.</b>
3.2. Software Configuration Management.....	<b>Error! Bookmark not defined.</b>

---

3.2.1 OCR Requirements .....	<b>Error! Bookmark not defined.</b>
3.2.2 Recognizing Process .....	<b>Error! Bookmark not defined.</b>
3.2.3 Verifying Process .....	<b>Error! Bookmark not defined.</b>
3.2.4 OCR for Mobile Devices .....	<b>Error! Bookmark not defined.</b>
3.2.4.1. Tesseract .....	<b>Error! Bookmark not defined.</b>
3.2.4.2. Abby Fine Reader OCR.....	<b>Error! Bookmark not defined.</b>
3.2.4.3. Omni page.....	<b>Error! Bookmark not defined.</b>
3.2.5 Comparison between different OCR for mobile devices	<b>Error! Bookmark not defined.</b>
3.3 Supplementary Specification .....	<b>Error! Bookmark not defined.</b>
3.3.1 Tesseract OCR .....	<b>Error! Bookmark not defined.</b>
3.3.2 Machine Translation .....	<b>Error! Bookmark not defined.</b>
3.3.2.1 MT process (MT stands for machine translation).....	<b>Error! Bookmark not defined.</b>
3.3.3 MT types .....	<b>Error! Bookmark not defined.</b>
3.3.3.1 Rule-Based Machine Translation.....	<b>Error! Bookmark not defined.</b>
3.3.3.2 Statistical Machine Translation.....	<b>Error! Bookmark not defined.</b>
3.3.3.3 Hybrid Machine Translation .....	<b>Error! Bookmark not defined.</b>
3.3.4 Challenges in Machine Translation.....	<b>Error! Bookmark not defined.</b>
3.3.4.1 Lexical Differences.....	<b>Error! Bookmark not defined.</b>
3.3.4.2 Syntactic Differences .....	<b>Error! Bookmark not defined.</b>
3.3.5 Online Machine translation .....	<b>Error! Bookmark not defined.</b>
3.3.5.1 Google Translator .....	<b>Error! Bookmark not defined.</b>
3.3.5.2 Bing Translator .....	<b>Error! Bookmark not defined.</b>
3.3.5.3 Babylon.....	<b>Error! Bookmark not defined.</b>
3.3.5.4 Word Lingo.....	<b>Error! Bookmark not defined.</b>
3.3.5.5 IM Translator .....	<b>Error! Bookmark not defined.</b>
3.3.6 Comparison between Online Machine translation.....	<b>Error! Bookmark not defined.</b>
3.3.7 Bing Translator.....	<b>Error! Bookmark not defined.</b>
3.3.8 Requirements Specification and Analysis .....	<b>Error! Bookmark not defined.</b>
3.3.8.1 Requirements Specification .....	<b>Error! Bookmark not defined.</b>
3.3.8.1.1 Requirements .....	<b>Error! Bookmark not defined.</b>

---

3.3.8.1.2 Chile Module .....	<b>Error! Bookmark not defined.</b>
3.3.8.2 Feasibility Report.....	<b>Error! Bookmark not defined.</b>
3.3.8.2.1 Technical Feasibility.....	<b>Error! Bookmark not defined.</b>
3.3.8.2.2 Client/Tourist.....	<b>Error! Bookmark not defined.</b>
3.3.8.2.3 Social Feasibility.....	<b>Error! Bookmark not defined.</b>
3.3.9 System Analysis.....	<b>Error! Bookmark not defined.</b>
3.3.9.1 Operational Feasibility.....	<b>Error! Bookmark not defined.</b>
3.3.9.2 Economic Feasibility .....	<b>Error! Bookmark not defined.</b>
3.3.10 Features and Services.....	<b>Error! Bookmark not defined.</b>
3.3.10.1 Functional Requirements .....	<b>Error! Bookmark not defined.</b>
3.3.10.2 Non-Functional Requirement.....	<b>Error! Bookmark not defined.</b>
3.3.10.3 Performance Requirements.....	<b>Error! Bookmark not defined.</b>
3.4 Software Quality Assurance Plan .....	<b>Error! Bookmark not defined.</b>
3.4.1 Purpose.....	<b>Error! Bookmark not defined.</b>
3.4.2 Scope.....	<b>Error! Bookmark not defined.</b>
3.4.3 Reference Documents .....	<b>Error! Bookmark not defined.</b>
3.4.4 Management.....	<b>Error! Bookmark not defined.</b>
3.4.4.1 Management Organization.....	<b>Error! Bookmark not defined.</b>
3.4.4.2 Responsibilities .....	<b>Error! Bookmark not defined.</b>
3.4.4.3 Review/Evaluation of Software Products .....	<b>Error! Bookmark not defined.</b>
3.4.4.4 Evaluation of Software Tools .....	<b>Error! Bookmark not defined.</b>
3.4.4.5 Evaluation of Software Requirement Analysis Process.....	<b>Error! Bookmark not defined.</b>
3.4.4.6 Evaluate Software Design Process.....	<b>Error! Bookmark not defined.</b>
3.4.4.7 Evaluate Software Implementation and Unit Testing Process.....	<b>Error! Bookmark not defined.</b>
3.4.5 Documentation.....	<b>Error! Bookmark not defined.</b>
3.4.5.1 Purpose.....	<b>Error! Bookmark not defined.</b>
3.4.6 Minimum Documentation Requirements.....	<b>Error! Bookmark not defined.</b>
3.4.6.1 Project Plan .....	<b>Error! Bookmark not defined.</b>
3.4.6.2 Software Requirements Specification (SRS).....	<b>Error! Bookmark not defined.</b>

---

3.4.6.3 Design Document.....	<b>Error! Bookmark not defined.</b>
3.4.6.4 Testing analysis.....	<b>Error! Bookmark not defined.</b>
3.4.6.5 User Manual.....	<b>Error! Bookmark not defined.</b>
3.4.6.6 End Report .....	<b>Error! Bookmark not defined.</b>
3.4.7 Standard and Conventions .....	<b>Error! Bookmark not defined.</b>
3.4.7.1 Purpose.....	<b>Error! Bookmark not defined.</b>
3.4.7.2 Project Implementation Standards:.....	<b>Error! Bookmark not defined.</b>
3.4.7.3 Documentation Standards: .....	<b>Error! Bookmark not defined.</b>
3.4.7.4 Metrics and Conventions: .....	<b>Error! Bookmark not defined.</b>
3.4.8 Software Reviews .....	<b>Error! Bookmark not defined.</b>
3.4.8.1 Purpose.....	<b>Error! Bookmark not defined.</b>
3.4.8.2 Deliverable Reviews for Our Project.....	<b>Error! Bookmark not defined.</b>
3.4.8.3 Audits:.....	<b>Error! Bookmark not defined.</b>
3.4.9 Validation and Verification.....	<b>Error! Bookmark not defined.</b>
3.4.10 Test.....	<b>Error! Bookmark not defined.</b>
3.4.10.1 Unit Test.....	<b>Error! Bookmark not defined.</b>
3.4.10.2 Integration Test .....	<b>Error! Bookmark not defined.</b>
3.4.11 System Testing.....	<b>Error! Bookmark not defined.</b>
3.4.12 Problem Reporting and Correct Action .....	<b>Error! Bookmark not defined.</b>
3.5 Use Case Model .....	<b>Error! Bookmark not defined.</b>
3.6 Use Cases (Essential).....	<b>Error! Bookmark not defined.</b>
3.6.1Use Case (Elaborated).....	<b>Error! Bookmark not defined.</b>
01 Use Case: Invoke Camera.....	<b>Error! Bookmark not defined.</b>
02 Use Case: Selection ROI (Region of Interest).....	<b>Error! Bookmark not defined.</b>
03 Use Case: Capture Image.....	<b>Error! Bookmark not defined.</b>
04 Use Case: Connect to OCR and Send Image to OCR.....	<b>Error! Bookmark not defined.</b>
05 Use Case: Showed Recognized Text .....	<b>Error! Bookmark not defined.</b>
06 Use Case: Copy Recognized Text.....	<b>Error! Bookmark not defined.</b>
07 Use Case: Connect with a Bing translator and send this recognized text for translate	<b>Error!</b>
	<b>Bookmark not defined.</b>

---

08 Use Case: Show Translated Text .....	<b>Error! Bookmark not defined.</b>
09 Use Case: Copy Translated Text.....	<b>Error! Bookmark not defined.</b>
10 Use Case: Speak Translated Text .....	<b>Error! Bookmark not defined.</b>
3.7 Software Architecture .....	<b>Error! Bookmark not defined.</b>
3.7.1 Data Flow Diagrams .....	<b>Error! Bookmark not defined.</b>
3.7.1.1 OCR Translator DFD.....	<b>Error! Bookmark not defined.</b>
3.7.1.2 Level 0 Diagrams .....	<b>Error! Bookmark not defined.</b>
3.7.1.3 Level 1 Diagrams .....	<b>Error! Bookmark not defined.</b>
3.7.1.4 Level 2 Diagrams .....	<b>Error! Bookmark not defined.</b>
3.8 Software Architecture .....	<b>Error! Bookmark not defined.</b>
3.8.1 Introduction.....	<b>Error! Bookmark not defined.</b>
3.8.2 System Design .....	<b>Error! Bookmark not defined.</b>
3.8.3 State Transition Diagram .....	<b>Error! Bookmark not defined.</b>
3.8.3.1 State Transition Table .....	<b>Error! Bookmark not defined.</b>
3.8.4 Context Diagram.....	<b>Error! Bookmark not defined.</b>
3.8.5 Activity Diagram .....	<b>Error! Bookmark not defined.</b>
3.9 Screen Shots.....	<b>Error! Bookmark not defined.</b>
3.9.1 Recognition .....	<b>Error! Bookmark not defined.</b>
3.9.1.1 English Recognizing .....	<b>Error! Bookmark not defined.</b>
3.9.1.2 Arabic Recognizing .....	<b>Error! Bookmark not defined.</b>
3.9.1.3 Bulgarian Recognizing.....	<b>Error! Bookmark not defined.</b>
3.9.1.4 Catalan Recognizing .....	<b>Error! Bookmark not defined.</b>
3.9.2 Copy Recognized Text.....	<b>Error! Bookmark not defined.</b>
3.9.3 Translation .....	<b>Error! Bookmark not defined.</b>
3.9.3.1 Translation Setting.....	<b>Error! Bookmark not defined.</b>
3.9.3.2 Bing Translator .....	<b>Error! Bookmark not defined.</b>
3.9.4 Copy Translated Text.....	<b>Error! Bookmark not defined.</b>
3.9.5 TTS (Text To Speech) .....	<b>Error! Bookmark not defined.</b>
Chapter: 4.....	<b>Error! Bookmark not defined.</b>
4.DESIGN DOCUMENTS .....	<b>Error! Bookmark not defined.</b>

---

4.1.Domain Model .....	<b>Error! Bookmark not defined.</b>
4.2.System Sequence Diagrams .....	<b>Error! Bookmark not defined.</b>
4.2.1 Crop Image.....	<b>Error! Bookmark not defined.</b>
4.2.2 Recognition .....	<b>Error! Bookmark not defined.</b>
4.2.3 Translation .....	<b>Error! Bookmark not defined.</b>
4.2.4 TTS (Text to Speak).....	<b>Error! Bookmark not defined.</b>
4.3.Sequence/Collaboration Diagram .....	<b>Error! Bookmark not defined.</b>
4.3.1 Over All View of System.....	<b>Error! Bookmark not defined.</b>
4.3.1.1 Descriptions .....	<b>Error! Bookmark not defined.</b>
4.4 Step by Step Description.....	<b>Error! Bookmark not defined.</b>
4.4.1 Crop Image.....	<b>Error! Bookmark not defined.</b>
4.4.2 Recognition Text.....	<b>Error! Bookmark not defined.</b>
4.4.3 Translation Text.....	<b>Error! Bookmark not defined.</b>
4.4.4 TTS (Text To Speech) .....	<b>Error! Bookmark not defined.</b>
4.5 Design Class Diagram.....	<b>Error! Bookmark not defined.</b>
4.5.1 Class Diagram .....	<b>Error! Bookmark not defined.</b>
4.5.2 Package Diagram .....	<b>Error! Bookmark not defined.</b>
4.6 Component Diagram.....	<b>Error! Bookmark not defined.</b>
4.6.1 Introduction and Description .....	<b>Error! Bookmark not defined.</b>
4.6.2 Diagram.....	<b>Error! Bookmark not defined.</b>
4.7 Deployment Diagram.....	<b>Error! Bookmark not defined.</b>
Chapter 5 .....	<b>Error! Bookmark not defined.</b>
5. TESTING.....	<b>Error! Bookmark not defined.</b>
5.1 Test Case Specifications .....	<b>Error! Bookmark not defined.</b>
5.1.1 Psychology of Testing.....	<b>Error! Bookmark not defined.</b>
5.1.2 Testing Objectives .....	<b>Error! Bookmark not defined.</b>
5.1.3 Levels of Testing.....	<b>Error! Bookmark not defined.</b>
5.1.4 System Testing.....	<b>Error! Bookmark not defined.</b>
5.1.5 Code Testing .....	<b>Error! Bookmark not defined.</b>
5.2Black Box Test Cases .....	<b>Error! Bookmark not defined.</b>

---

5.2.1 Introduction.....	<b>Error! Bookmark not defined.</b>
5.2.2 Test Cases .....	<b>Error! Bookmark not defined.</b>
5.3 White Box Test Cases.....	<b>Error! Bookmark not defined.</b>
5.3.1 Cyclamate Complexity.....	<b>Error! Bookmark not defined.</b>
5.3.2 Cyclamate complexity and Control Flow Graph .....	<b>Error! Bookmark not defined.</b>
5.4 Performance Testing.....	<b>Error! Bookmark not defined.</b>
5.4.1 Load Testing: .....	<b>Error! Bookmark not defined.</b>
5.4.2 Stress testing .....	<b>Error! Bookmark not defined.</b>
5.5 System Testing.....	<b>Error! Bookmark not defined.</b>
5.5.1 Types of System Testing.....	<b>Error! Bookmark not defined.</b>
5.6 Regression Testing.....	<b>Error! Bookmark not defined.</b>
Chapter 6.....	<b>Error! Bookmark not defined.</b>
6. TOOLS USED .....	<b>Error! Bookmark not defined.</b>
6.1 Languages .....	<b>Error! Bookmark not defined.</b>
6.1.1 Java Language.....	<b>Error! Bookmark not defined.</b>
6.1.2 XML Language.....	<b>Error! Bookmark not defined.</b>
6.2 Applications .....	<b>Error! Bookmark not defined.</b>
6.2.1 Software Requirement .....	<b>Error! Bookmark not defined.</b>
6.2.2 Hardware Requirement .....	<b>Error! Bookmark not defined.</b>
6.2.3 Minimum Required Hardware .....	<b>Error! Bookmark not defined.</b>
6.2.4 Android .....	<b>Error! Bookmark not defined.</b>
6.2.5 Eclipse.....	<b>Error! Bookmark not defined.</b>
Chapter 7.....	<b>Error! Bookmark not defined.</b>
7.SUMMARY AND CONCLUSIONS .....	<b>Error! Bookmark not defined.</b>
7.1 Summary.....	<b>Error! Bookmark not defined.</b>
7.2 Conclusion .....	<b>Error! Bookmark not defined.</b>
Chapter: 8.....	<b>Error! Bookmark not defined.</b>
8.LESSON LEARNT AND FUTURE WORK .....	<b>Error! Bookmark not defined.</b>
8.1 Lesson Learnt.....	<b>Error! Bookmark not defined.</b>
8.2 Future Work.....	<b>Error! Bookmark not defined.</b>

---

Appendix A.....	<b>Error! Bookmark not defined.</b>
Project Management Documents:.....	<b>Error! Bookmark not defined.</b>
A) Gantt chart .....	<b>Error! Bookmark not defined.</b>
b) Pert Diagram.....	<b>Error! Bookmark not defined.</b>
Appendix B .....	<b>Error! Bookmark not defined.</b>
References/Bibliography.....	<b>Error! Bookmark not defined.</b>

---

## TABLE OF FIGURES

---

Figure 3.4 Use Case Diagram .....	<b>Error! Bookmark not defined.</b>
Figure 3.7.1 Data Flow Diagram .....	<b>Error! Bookmark not defined.</b>
Figure 3.8.2 System Design .....	<b>Error! Bookmark not defined.</b>
Figure 3.8.3 State Transition Diagram.....	<b>Error! Bookmark not defined.</b>
Figure 3.8.4 Context Diagram .....	<b>Error! Bookmark not defined.</b>
Figure 3.8.5 Activity Diagram .....	<b>Error! Bookmark not defined.</b>

---

Figure 3.9.1 Capturing Image .....	<b>Error! Bookmark not defined.</b>
Figure 3.9.2 Capturing Image Using ROI.....	<b>Error! Bookmark not defined.</b>
Figure 3.9.3 Recognizing Language .....	<b>Error! Bookmark not defined.</b>
Figure 3.9.4 English recognizing .....	<b>Error! Bookmark not defined.</b>
Figure 3.9.5 Arabic recognizing .....	<b>Error! Bookmark not defined.</b>
Figure 3.9.5 Arabic recognizing .....	<b>Error! Bookmark not defined.</b>
Figure 3.9.7 Catalan recognizing .....	<b>Error! Bookmark not defined.</b>
Figure 3.9.8 Copy Recognizing Text.....	<b>Error! Bookmark not defined.</b>
Figure 3.9.9 Translation Setting.....	<b>Error! Bookmark not defined.</b>
Figure 3.9.10 Translation Language .....	<b>Error! Bookmark not defined.</b>
Figure 3.9.11 Translator.....	<b>Error! Bookmark not defined.</b>
Figure 3.9.12 Translations .....	<b>Error! Bookmark not defined.</b>
Figure 3.9.13 Popup for Copy Translated Text .....	<b>Error! Bookmark not defined.</b>
Figure 3.9.13 Popup for Speak Translated Text .....	<b>Error! Bookmark not defined.</b>
Figure 3.9.14 Press to Speech Button for that Translated Text ...	<b>Error! Bookmark not defined.</b>
Figure 4.1 Domain Model.....	<b>Error! Bookmark not defined.</b>
Figure Crop Image (System Sequence Diagram) .....	<b>Error! Bookmark not defined.</b>
Figure Recognition.....	<b>Error! Bookmark not defined.</b>
Figure Translation.....	<b>Error! Bookmark not defined.</b>
Figure Text to Speech .....	<b>Error! Bookmark not defined.</b>
Figure Sequence diagram.....	<b>Error! Bookmark not defined.</b>
Figure Crop image .....	<b>Error! Bookmark not defined.</b>
Figure Recognition text.....	<b>Error! Bookmark not defined.</b>
Figure Translation text.....	<b>Error! Bookmark not defined.</b>
Figure Text to Speech .....	<b>Error! Bookmark not defined.</b>
Figure Class Diagram .....	<b>Error! Bookmark not defined.</b>
Figure Component Diagram .....	<b>Error! Bookmark not defined.</b>
Figure Deployment Diagram .....	<b>Error! Bookmark not defined.</b>
Figure 5.5.1 Type of Testing.....	<b>Error! Bookmark not defined.</b>
Figure Pert Diagram.....	<b>Error! Bookmark not defined.</b>

---

## LIST OF TABLES

---

Table 1 Comparison between Online Machine Translation	<b>Error! Bookmark not defined.</b>
Table 2 State Transition Table	<b>Error! Bookmark not defined.</b>

---

Table 3 Software Requirements

Table 4 Minimum Hardware Requirements

**Error! Bookmark not defined.**

**Error! Bookmark not defined.**

---