ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 3, Issue 2, March 2014

An Optimal Text Recognition and Translation System for Smart phones Using Genetic Programming and Cloud

Ashish Emmanuel S, Dr. S.Nithyanandam

Abstract— An Optimal Text Recognition and Translation System for Smart phones using Genetic Programming and cloud processes image with genetic algorithms is a live video translator which translates spontaneously. It uses optical character recognition (OCR) to recognize the characters and can be really useful when you have got a paper document you want in digital, editable form. You need a scanner which can be your mobile phone to create an image of the document first, but then once you have that image you can run it through this application that will read each character and attempt to recreate the original document as text. The image, even a road sign, is first preprocessed by using image acquisition process. Then the image segmentation is performed. The image is then processed with genetic algorithms, which is used to process the image to maximize the quality for text conversion from image. Preprocessed image is asynchronously transferred to character identification web services and obtained as actual text. Once the data is processed, it is displayed to the user. This system can translate any text written in English to Tamil also through translating web services.

Index Terms **- Camera-based Translation, OCR with Cloud, Genetic Programming, Computer Vision, Text Recognition.**

I. INTRODUCTION

Optimal Text Recognition and Translation System for Smart phones using Genetic Programming & Cloud can automatically detect the language and translate it spontaneously. Any language can be converted into any other language. It is very useful for people who visit various countries of unknown language because they can read signs and boards with the use of this application. We can take a picture of any document and this application turns it into text and searches web for more information. This can also be sent as SMS, mail or anything else.

This system uses the Genetic algorithm which is used to process the image and produce maximum image quality by removing the noise and separate text from its background. The image segmentation is also performed and the color text is converted into grayscale and then the resultant image is translated by obtaining data from translation services. Then any massive changes in the text are noticed and if there are any changes present, the text is again sent to translation services and the obtained text is displayed as output. Though multiple attempts encounter traffic, it doesn't exceed the limit as in the previous system since the image is preprocessed to reduce its size and produce better and accurate results.

Translated can be done easily by looking at the street sign using our smart phone's camera. It auto-focuses your camera as you point (if auto-focus is supported by your phone). While taking images, you can either snap a full image or highlight the area you want to translate with a rectangular box. Phone's flash can also turn on automatically based on lighting conditions. This also can save the text in the phone's clipboard to send as text message, mail, use in any app or to search in the web. This system can translate any text written in English to Tamil also through translating web services.

II. EXISTING SYSTEM

In existing system they have developed a prototype of a mobile app called "Snap and Translate" on "Windows Phone 7".In that if a person who is reading an English menu/sign and wants a Chinese translation of an English word or phrase or paragraph can use a Windows Phone to snap an image of the text, tap the word or swipe the phrase or circle the paragraph with a finger, and get a Chinese translation displayed on the screen of the phone [1]. They enable this translation by seamless integration of three Microsoft technologies: intelligent text extraction, OCR, and machine translation based on a client-plus-cloud architecture [1]. In existing system they highlight the

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 3, Issue 2, March 2014

UI design of the system and the corresponding user-intention guided text extraction approach to achieving a compelling user experience. The UI that uses the swipe gesture is affected by allowing the user to make vertical corrective movements first to decide the starting target point, followed by horizontal corrective movements until the pointer is over the ending target point to lift the finger for target selection. The third gesture is *circle* that allows the user to circle the intended text area, which may contain a paragraph with multiple text lines.

III. PROPOSED SYSTEM

In the proposed system, the image is first preprocessed by using image acquisition process as the live video text will not be in proper lighting and background conditions. For acquiring a high quality image, the various types of filters are used. The image acquisition helps enhancing the image. It also performs image smoothing and converts the color image to gray scale for identifying the text properly. Then the image segmentation is performed and the image is divided into parts for processing the image quickly. This is mainly performed to separate the text from the background. Then, with the help of the genetic algorithms [2], the quality of the image is improved thereby making it easier for the text identification algorithms. But the existing system uses the edge based text detection algorithm which recognizes text with gestures such as drawing a box, tap or circle. The size of processed image is reduced massively and the image segmentation helps in reducing it further thus letting to use multiple asynchronous attempts for translating text [3]. In our application, maximum effort is made to identify the text from the image and to produce more accurate results than the existing systems.

IV. IMAGE PREPROCESSING

In this system, the image is first preprocessed by using image acquisition process as the live video text will not be in proper lighting and background conditions. The image acquisition helps enhancing the image. Then the image segmentation is performed and the color text is converted into gray scale after which the image is processed with genetic algorithms to improve the quality of the image thereby making it easier for the text identification algorithms. But the existing systems use the edge based text detection algorithm which recognizes text with gestures such as drawing a box, tap or circle. The size of processed image is reduced massively and the image segmentation helps in reducing it further thus letting to use multiple asynchronous attempts for translating text. In our application, maximum effort is made to identify the text from the image and to produce more accurate results than the existing systems.

V. IMAGE ACQUISITION

Image Acquisition is the first stage of any image based analysis process. In this module, the communication with the mobile camera takes place as live camera image is used for processing. As it is a live camera image, it will not be in a perfect lighting and background. So, image is preprocessed here to enhance the image. The color image is converted to a gray scale image and image smoothing and noise removal are performed.

A major factor involved in image acquisition in image processing sometimes is the initial setup and long-term maintenance of the hardware used to capture the images. The actual hardware device used here is the windows phone. So, improper setup of hardware problem won't occur and also it will not provide images that are of such low quality that they cannot be salvaged even with extensive processing. One of the forms of image acquisition in image processing is known as real-time image acquisition. This usually involves retrieving images from a source that is automatically capturing images.

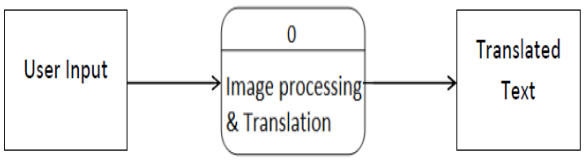

Fig.1 Level 0 DFD

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 3, Issue 2, March 2014

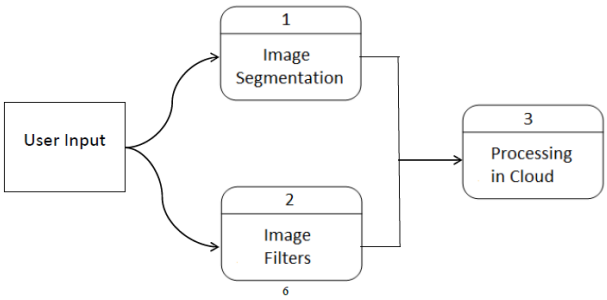

Fig.2 Level 1 DFD

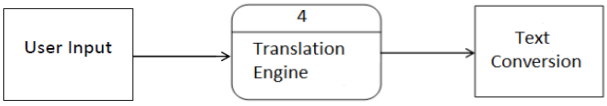

Fig.3 Level 2 DFD

VI. IMAGE SEGMENTATION

The goal here is to cut out the text from its background. Usually a small size image is preferable for processing as the processing power is limited in mobiles. So, the image size is reduced to a great extent by allowing maximum thresholds.

Segmentation techniques are either contextual or non-contextual. The latter take no account of spatial relationships between features in an image and group pixels together on the basis of some global attribute, e.g. grey level or color. Contextual techniques additionally exploit these relationships, e.g. group together pixels with similar grey levels and close spatial locations. Segmentation could be used for object recognition, image compression, image editing, etc. But here we use segmentation to separate the text from the background. The quality of segmentation depends on the image. Smoothly shaded surfaces with clear gray scale level are ideal for segmentation.

VII. GENETIC ALGORITHM

This is an evolutionary algorithm-based methodology inspired by biological evolution to find computer programs that perform a user-defined task. Here this algorithm is used to process the Image to maximize the Quality for Text Conversion from Image. Experiments on images from the Berkeley Image Segmentation Dataset show that the method is able to partition natural and human scenes in a number of regions consistent with human visual perception. A quantitative evaluation of the method compared with other approaches shows that the genetic algorithm can be very competitive in partitioning color images [2].

VIII. TEXT CONVERSION

Preprocessed Image is asynchronously transferred to character identification web services and obtained as actual text. This Converted Text is sent for translation services asynchronously. The exact mechanisms that allow humans to recognize objects are yet to be understood, but the three basic principles are already well known by scientists – integrity, purposefulness and adaptability (IPA). These principles constitute the core of OCR allowing it to replicate natural or human-like recognition [2].

IX. TRANSLATION ENGINE

Once the data is processed, it is displayed to the user [1]. At the same time, another thread performs the same text identification and finds the changes in the text; if any massive changes exist in the converted text, another attempt of the same will be done by sending the newly obtained data to the translation service. Though multiple attempts

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 3, Issue 2, March 2014

encounter traffic, it doesn't exceed the limit as in the previous system since the image is preprocessed to reduce its size and produce better and accurate results.

X. ADVANTAGES OF OPTIMAL TEXT RECOGNITION & TRANSLATION SYSTEM

- The text to be translated is automatically detected.
- As live camera stream image is considered for processing, the processing time is reduced.
- The size of the image is also reduced for a greater extent during preprocessing itself which reduces the complexity during translation of text.
- This mechanism offers more accurate results.

XI. COMPARISON WITH EXISTING SYSTEM

In existing systems it uses the edge based text detection algorithm for recognizing the text with gestures such as drawing a box, tap or circle. But in our proposed system the text image can be extracted automatically on the phone and send to the cloud for recognition and translation. Therefore this text translation with preprocessing filters results will be more accurate and needs only less computation and also it uses only less network traffic as shown in Fig.4 and Fig.5.

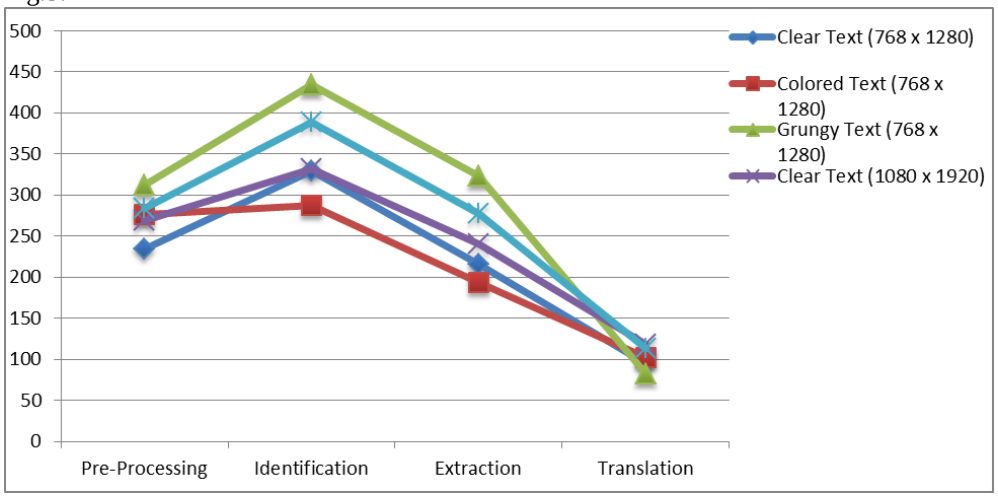

Fig.4 Time Taken For Each Stage

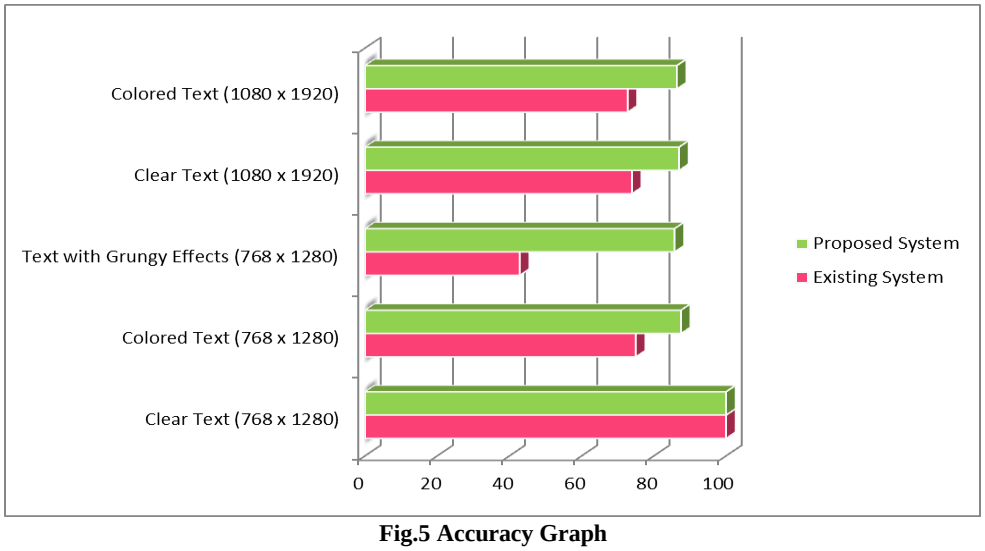

Fig.5 Accuracy Graph

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 3, Issue 2, March 2014

XII. CONCLUSION

An Optimal Text Recognition & Translation System for Smart phones using Genetic Programming & Cloud processes image with genetic algorithms to improve the quality of the image thereby making it easier for the text identification.

The processed image size is also reduced to a great extent which reduces the segmentation and allows multiple asynchronous attempts. It works well for scenarios such as translation of text in printed newspapers, magazines, books, menus, and signboards with simple backgrounds under any lighting conditions.

It produces more accurate results than the already existing application by using pre-processing filters as shown in Fig.6, Fig.7 and Fig.8. The improved text extraction and OCR engine in used to deal with more challenging scenarios such as text with complex backgrounds and uncommon font types under non-uniform lighting conditions and casual image capturing with large geometry distortions as in Fig.9 and Fig.10.

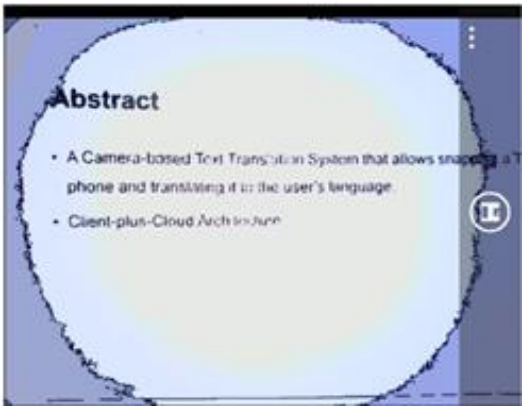

Fig.6 Edge-Detection

Fig.7 Applying RBG Filter based on Color Adjustment

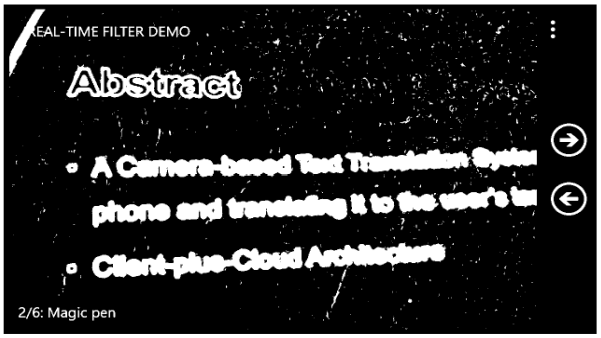

Fig.8 Applying Threshold Filter (which is a Combination of Grayscale and Contrast Filter)

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 3, Issue 2, March 2014

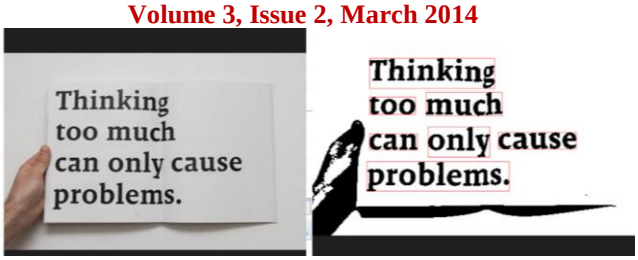

Fig.9 Text Identification from Image and highlighting it

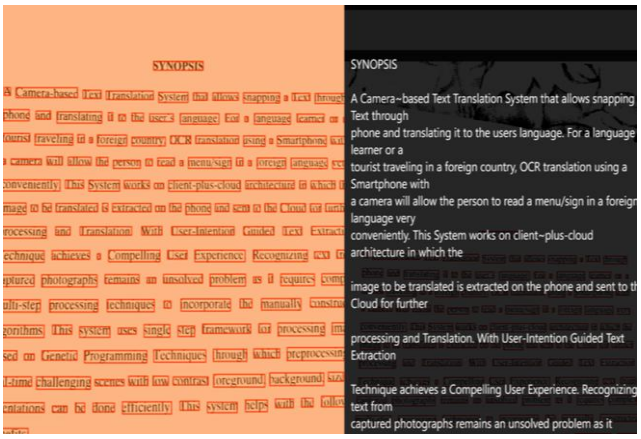

Fig.10 Text Extraction from the Identified Text

Fig.11 Translating the Identified Text to the User Language

ACKNOWLEDGMENT

First and foremost, I would like to thank God Almighty and My Beloved Parents without whom I'll be unable to complete my project on time. I deemed it as a pleasure and privilege to feel indebted to my guide Dr. S. Nithyanandam, M.Tech., PhD., Associate Professor, Department of Computer Science and Engineering, PRIST University, Thanjavur, for his innovative suggestions and valuable hints in the successful completion of my project. I would like to convey my sincere thanks to all the Faculties, Department of Computer Science & Engineering, PRIST University, Thanjavur, Without their cooperation, this venture would not have been a success.

REFERENCES

[1] Snap and Translate Using Windows Phone (IEEE.2011), Jun Du, Qiang Huo, Lei Sun, Jian Sun Microsoft Research Asia, Beijing, P. R. China.

[2] Towards Scene Text Recognition with Genetic Programming (IEEE.June,2013) , Brendan Barlow, Andy Song School of Computer Science and IT, RMIT University, Australia.

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 3, Issue 2, March 2014

[3] V. Fragoso, S. Gauglitz, S. Zamora, J. Kleban, M. Turk "TranslatAR: a mobile augmented reality translator," Proc. IEEE Workshop on Applications of Computer Vision (WACV) 2011.

[4] Google Goggles, http://www.google.com/mobile/goggles/#text

[5] J. Yang, J. Gao, Y. Zhang, A. Waibel, "Towards automatic sign translation," Proc. HLT-2001.

[6] I. Haritaoglu, "Scene text extraction and translation for handheld devices," Proc. CVPR-2001, pp.II-408-413.

[7] D. Chen, J.-M. Odobez, and H. Bourlard. Text detection and recognition in images and video frames. Pattern Recognition,37(3):595 – 608, 2004.

AUTHOR BIOGRAPHY

Ashish Emmanuel S. is doing his M.Tech degree in Computer Science and Engineering in PRIST University. He received his B.E. degree from Anna University, Chennai in 2012. He has attended two International conferences. He has worked in a lot of innovative real-time projects on Computer Vision, Gesture Recognition, Gesture User Interfaces, Mobile Application Development and Networking during the course of his degree completion.

Nithyanandam.S is an Associate Professor in the CSE Department. He received P.hD Degree in Computer Science & Engg. in PRIST University, He got his M.Tech Degree in Information from Punjabi University Technology, Patiala in 2003, and also received MCA.,M.Sc(Physics), & M.Phil. Degrees from Bharathidasan University, Tiruchirappalli. He has presented 6 Papers in international / National Conferences and published 8 Papers in reputed International Journals. He also published 3 books namely Computer Networks, Mobile Computing & System Software Published by SAMS Publishers and Dhanpatrai & Co,New Delhi.