PRINTED BOOK CONVERTED INTO AUDIO FORMAT

[1] Rushikesh.S.Maratkar, [2] K.S Karpe, [3]Shreyash Pathak, [4]Narayan Patil
[1,3,4] Student, [2] Asst prof. Depatment of electronics and telecommunication PVPIT,
Savitribai phule Pune University, Pune , Maharashtra, India.
Email: [1] rushikesh.maratakar1993@gmail.com

Abstract

Optical Character recognition is a process of converting printed or scanned documents into ASCII characters that a computer can recognize. Character recognition is one of the important tasks in Pattern Recognition. The complexity of the character recognition problem depends on the character set to be recognized. The Proposed system present to develop an optical character recognition technique for English text images to convert them into the audio. The output of the Text to Speech System Optical Character recognition system is the audio corresponding to text in input image. Here, two prime techniques Optical Character recognition and application program interface are used to convert the input text image into audio.

Keywords*:* OCR, Binarization, thresholding, segmentation, template creation & matching, windows API.

I. INTRODUCTION

Speech is probably the most efficient medium for communication between humans. Optical character recognition has become one of the most successful applications of technology in the field of pattern recognition and artificial intelligence. Character recognition or optical character recognition , is the process of converting scanned images of machine printed or handwritten text like numerals, letters, and symbols into a computer format text. Speech synthesis is the artificial synthesis of human speech. A Text-To-Speech synthesizer is a computer-based system that should be able to read any text aloud, whether it was directly introduced in the computer by an operator or scanned and submitted to an Optical Character Recognition system. Operational stages of the system consist of image capture, image preprocessing, image filtering, character recognition and text to speech conversion. The software platforms used is MatLab.

II. A BRIEF MOTIVATION

The purpose of project is to make available any English text converted into audio book. For many blind users educational choices are made based on which material can be accessed. These people are dependent solely on Braille books & audio recordings provided by NGOs.The presented work will provide visually impaired people, an opportunity to have an audio material of their own choice of any printed material. The framework consists of two parts. One is Optical Character Recognition which includes operations like grayscaling, thresholding, filtering, thinning, segmentation, cropping, etc. on a character in the image and other part is text to speech conversion using Microsoft's API which will convert the text into speech (audio).

III. RELEVANCE

There are many documents in world which are not available in computer comprehensible format i.e. computer cannot read them. Computer vision and Robotics cannot be imagining without OCR. It acts as Benefit to Blind Persons. OCR has enabled scanned documents to become more than just image files, turning into fully searchable documents with text content that is recognized by computers. With the help of OCR, people no longer need to manually retype important documents when entering them into electronic databases. Instead, OCR extracts relevant information and enters it automatically. The result is accurate, efficient information processing in less time.

IV. SYSTEM BLOCK DIAGRAM

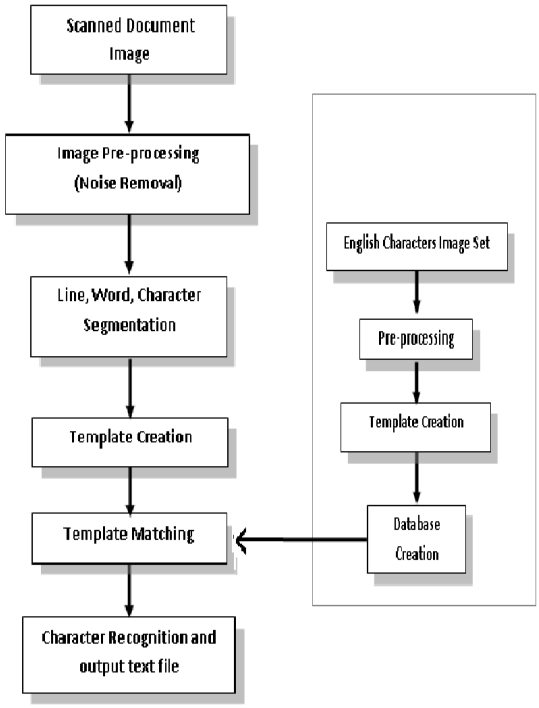

Figure 1 - OCR Block Diagram.

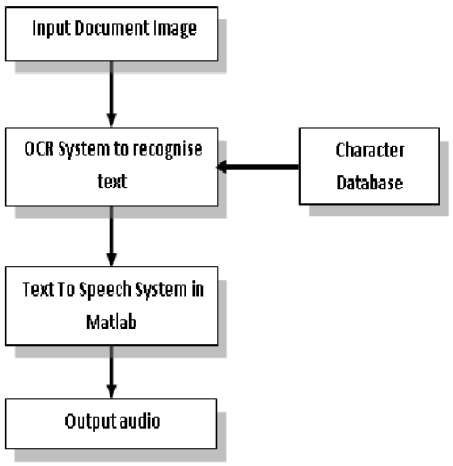

Fig 2 – Text to speech conversion

A. Input Scanned Image

By using camera we get the input printed document scanned and the output of camera will be image of printed book. Firstly image of input data is optically scanned. The scanned image can be any document of different dimensions. This scanned input image is fed to pre-processing section so as to process over that scanned image.

B. Pre-Processing –

Pre-processing includes several operations over the scanned image, so that input image becomes suitable and comfortable for applying to further sub sections. Basically the objective of pre-processing is to improve the quality of scanned input image. Noise removal, mathematical operations can also be processed in this Pre-processing section. It includes binarization, boundary detection, segmentation, thinning. It performs the several operations over the scanned input data.

Different morphological operations are performed in pre -processing to remove noise, watermarks ,and background contents in the image. The different morphological operations are –

1) Image resize
2) Beware opening
3) Imerode

C. Binarization

Binarization plays an important role in pre-processing. It is necessary to convert a color image into black and white format. So we can process over that black and white image. Basically separation of background and actual image area referred as foreground of a scanned image is called binarization.

D. Boundary Detection

The binarized image is now applicable for boundary detection. In this operation the boundaries of scanned image is detected. It detects all the boundaries of image. It is necessary to detect the boundaries so as to select an individual character.

E. Segmentation

This is important operation of OCR as rate of recognition is directly proportional to segmentation. In this process, firstly the every individual line of the image is separated and then each individual character of that line is separated. this process will be continued until the each and every line and character of the image will be separated. This isolates the different sub-parts of an image. It is used to separate pixels of an image as per the contents in data like words, paragraph etc.

F. TEMPLATE CREATION

In this process we creating data base of alphabets and numbers. In English there are total 26

alphabets and 0-9 that is 10 numbers. For this we are creating database that is templates. There are total 36 template which is of alpha bets and numbers. After creating templates these templates are then converted to cell format of size 42*24.
And saved as matrix.

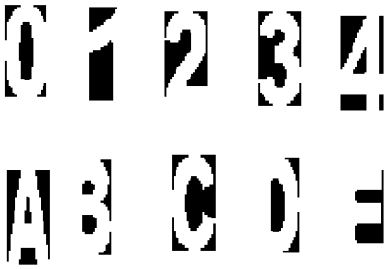

Fig 3 – template database.

G . TEMPLATE MATCHING

After creating these templates, they are matched with input scanned document. That is with the input image.after matching it is converted into text document. That is nothing but input printed document is converted to text file.

I. Text to Speech conversion

Using speech application program interface (SAPI) the given text is converted into audio. The Speech Application Programming Interface is an API developed by Microsoft to allow the use of speech recognition and speech synthesis within Windows applications. To date, a number of versions of the API have been released, which have shipped either as part of a Speech SDK, or as part of the Windows OS itself. Applications that use SAPI include Microsoft Office, Microsoft Agent and Microsoft Speech Server.

V. RESULTS

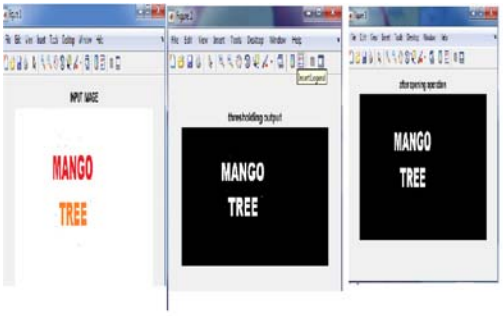

Fig 4 – Input image, After thresholding, After opening

Fig 5 – Line segmentation and character segmentation

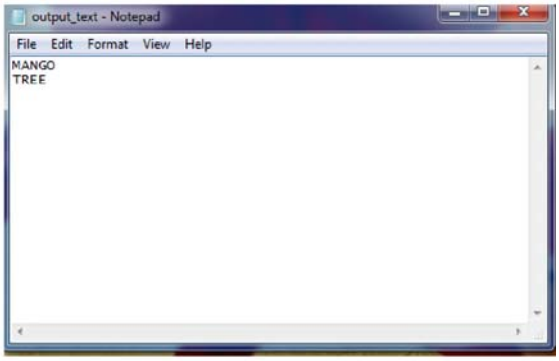

Fig 6 – Text file output of OCR

VI. OCR PERFORMANCE EVALUATION

No standardized test sets exist for character recognition, and as the performance of an OCR system is highly dependent on the quality of the input, this makes it difficult to evaluate and compare different systems. Still, recognition rates are often given, and usually presented as the percentage of characters correctly classified. However, this does not say anything about the errors committed. Therefore in evaluation of OCR system, three different performance rates are investigated:

Recognition rate

The proportion of correctly classified characters

Rejection rates

The proportion of characters which the system was unable to recognize.
Rejected characters can be flagged by the OCR-system, and are therefore easily retraceable for manual correction

Error rate

In this proposed system input image font must be Arial Black because the font of templates we have created is in Arial Black. That is nothing but database which we have created in particular font it should matched with input image.

Therefore input image font must be same as database font. Otherwise detection of the letters will not be proper and there are chances of faulty detection. And it introduces error.

VI. CONCLUSION

This project is an effort to suggest an approach for image to speech conversion using optical character recognition and text to speech technology. The application developed is user friendly, cost effective and applicable in the real time. By this approach we can read text from a document, Web page or e-Book and can generate synthesized speech through a computer's speakers or phone's speaker.

The developed software has set all policies of the singles corresponding to each and every alphabet, its pronunciation methodology, the way it is used in grammar and dictionary. This can save time by allowing the user to listen background materials while performing other tasks. System can also be used to make information browsing for people who do not have the ability to read or write. People with poor vision or visual dyslexia or totally blindness can use this approach for reading the documents and books.

The project presents complete Optical Character Recognition system followed by text to speech conversion. The basics of image processing and MATLAB are taken into consideration. Because of the computing constraints of handheld devices, study is being kept limited to light-weight and computationally efficient techniques. Various algorithms for optical character recognition have been studied. Modules related to OCR are implemented which determines the efficiency and accuracy of project.

VII. REFERENCES

[1] Jagruti Chandarana, Mayank Kapadia Optical Character Recognition International Journal of Emerging Technology and Advanced Engineering Volume 4, Issue 5, May 2014.

[2] Binod Kumar Prasad, GoutamSanyalA model Approach to Off-line English Character Recognition International Journal of Scientific and Research Publications, Volume 2, Issue 6, June 2012.

[3] SandeepTiwari, Shivangi Mishra, Priyank Bhatia, Praveen Km. Yadav[May 2013] Optical CharacterRecognition using MATLAB International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE)Volume 2, Issue 5, May 2013.

[4] Majida Ali Abed Hamid Ali Abed Alasadi Simplifying Handwritten Characters Recognition Using a Particle Swarm Optimization Approach European Academic Research, Volume I, Issue 5/ August 2013.

[5] Mahesh Goyani, HarshDani, Chahna Dixit Handwritten Character Recognition – A Comprehensive Review International Journal of Research in Computer.

[6] Gonzalez Rafael .C "Digital Image Processing" Pearson Education Second Edition, Upper Saddle River New Jersey, USA, 2002

[7] Chapman Stephen .J "MATLAB programming for Engineering" second edition, 2002

[8] Forsyth David .A, Ponce Jean "Computer Vision-A Modern Approach" Pearson Education, First Edition, Upper Saddle River New Jersey, USA, 2003

[9] Sandeep Tiwari, Shivangi Mishra, Priyank Bhatia, Praveen Km. Yadav, "Optical Character Recognition using MATLAB" International Journal of Advanced Research (2013), Volume 1, Issue 9, 757-767

[10] Oi-Mean Foong and Nurul Safwanah Bt Mohd Razali, "Signage Recognition Framework for Visually Impaired People", In.proc.of International Conference on Computer Communication and Management, Vol.5, 2011.