Scene To Text Conversion For Visually Impaired People

B.Anuradha[1], K.Harish[2], V.Dineshkumar[3], R.Dhanvish[4], S.Ganeshbharathi[5]
[1]Assistant Professor, Dept of Electronics and Communication
[2, 3, 4, 5]Dept of Electronics and Communication
[1, 2, 3, 4, 5] Sri Eshwar College Of Engineering, Coimbatore, TamilNadu.

Abstract- *The recent technological advancements are focusing on developing smart systems to improve the quality of life. Machine learning algorithms and artificial intelligence are becoming elementary tools, which are used in the establishment of modern smart systems across the globe. In this context, an effective approach is suggested for automated text detection and recognition for the natural scenes. The incoming image is firstly enhanced by employing Contrast Limited Adaptive Histogram Equalization (CLAHE). Afterward, the text regions of the enhanced image are detected by employing the Maximally Stable External Regions (MSER) feature detector. The non-text MSERs are removed by employing appropriate filters. The remaining MSERs are grouped into words. The text recognition is performed by employing an Optical Character Recognition (OCR) function. The extracted text is pronounced by using a suitable speech synthesizer. The proposed system prototype is realized. Results prove the concept and working principle of the devised system.*

Keywords- OCR, Text extraction, Segmentation, TTS Template, MATLAB16

I. INTRODUCTION

In the fastest moving world, new technology arises and more advancement is taking place day by day to make a smart system that improves the quality of life.From this project an innovative, real-time application which is cost-effective for visually impaired people. Who face many problems in the living world? It enables them to hear the content in the document /texted image. It was developed using the OCR technique which extracts text and pre-processes to recognize every letter. Then segment them to re-sizethose letters into a file containing the texts. At last, the text information is being converted into an audio signal. This process is taking place in MATLAB16

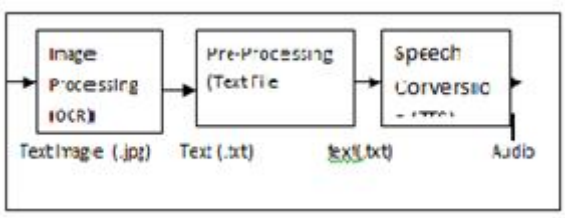

Figure.1: Scene to the speech Conversion process

II. EXISTING WORK

In 2017,IEEE transactions,Kumara et al [1] the paper describes the implementation of text is extracted and it converted into speech using Raspberry Pi.This Algorithm used is Complex. In 2016 on IEEE paper was published.Chandran et al[3] the paper describes the implementation of speech generation system using MATLAB, win 32 SAPI software (library) and creates a database. It takes more time to Stimulate.In 2018 on JOSCIENCE was published. Gopinath et al [4] the paper describes the image is converted into text then the text is converted into speech using MATLAB, lab view and android platform.Memory capacity is higher and does not work on all models.

III. PROPOSED WORK

The proposed system of a block diagram is shown in Figure2. The detailed description is explained below.

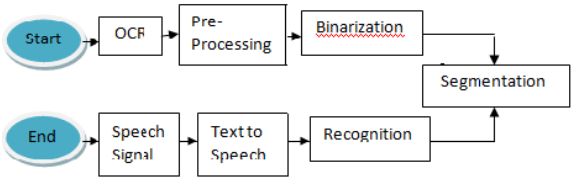

Figure2:Block Diagram1

1. Optical Character Recognition:

OCR is a reader that recognizes the text character of the computer which may be in printed or written form. The OCR templates of each character are used to recognize the character i.e. scanning process is carried out. After this, the

character image is translated in ASCII code which is further used in the data processing. The image recognition where each character from the image are being sent for recognition. There is some pre-processing stage involved to make the image noise free which involves process like binarization.

Pre-processing stage:

The Pre-processing stage eliminates all the challenges created by noise, and other additional errors. In this stage, the image inputis processed to remove any noise that may affect the image during the time of acquisition or during the time of transmission. A (RGB) image is converted to a grayscale image and then threshold, edge thinning, and noise removal process will be done. The image is again converted to a binary image with a suitable threshold to simplify the extraction process

(i)Grayscale
(ii) Binary image

The Original image (RGB)captured by the camera, grayscale image used for further process of extraction, and it is a binary image that computers can recognize. This process is a very important step in extracting text from the image, since RGB image may have noise, and could not identify text and non-text objects of the image.

Processing Stage:

The processing stage consists of various steps in which the image is checked whether it has text or not, which identifies to locate the text area on the image, and differentiating foreground and background of that image text are done.

2. Binarization:

A process of binarization is converting a grayscale image into a binary image using thresholding is known as Binarization. Before making the phenomenon acknowledged taking you the decades back, it was used in faxes now the binarization is easy but typical to understand in simple words we know that the image contains pixels which are stored bit by bit, now in the image there are two colors black(0)and white(1) what does it do is that grey pixels it makes them set(accepted) and the pixels with white are made unset further in the process the pixels which are inset mode and are near are combined to make some acceptable character. The important characteristic of the binarization is the distance transformation by which the unset pixel is distanced from another set pixel.

3. Segmentation:

Segmentation is done as the image consists of several sentences and each line contains a certain number of the words. Then this word is formed by the number of characters. Hence we can say that the segmentation is a process of partitioning the digital image into the segments. n the segmentation process each line, each word, each character is segmented. There are some inevitable difficulties in the segmentation process like image quality is less, every computer system has some different fonts, cursive writing, etc. this affects the efficiency of the segmentation process thus the best practice is done here to eliminate those problems.

4. Recognition:

Ones the character is segmented it is stored in a variable as to compare it with the stored template form. Precisely preliminary data will be stored in the form of templates where all characters recognized font and size are available. The data contains further information: the value of ASCII character, the name of the character, character of the image in Jpg and character width, etc. for every identified character all information will be captured then there is the comparison between the recognized character information with the predefined character which is stored in the system.

5. Text to Speech Conversion:

Finally, the artificial production of human sound is nothing but the speech. The speech generated by the mechanism is known as a speech synthesizer. After all process text images are being converted into speech signal, it consists of two ways front-end and back-end process. The front-end process is the extraction of the file, raw images, and processing them in the form of a single character and number. The back-end process is extracted text to audio or speech signal.

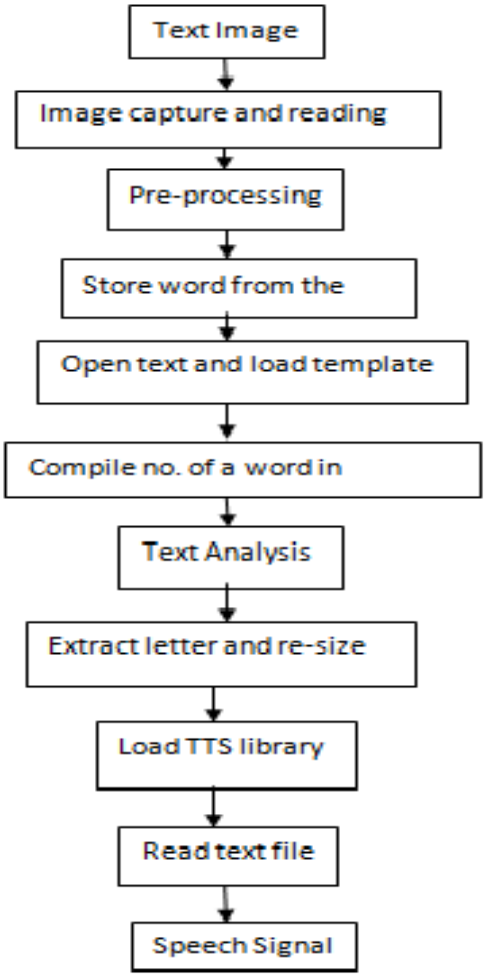

Figure 3 - Block Diagram2

Steps that take place in the above block diagram.

Step1: Image is being captured using web cam are the image files which are being stored.
Step2: The stored file is taken and in which the word are reader by OCR technique.
Step3: After that pre-processing take place in which the unwanted noise signal are being removed.
Step4: The alphabetic letters are being stored in separate.
Step5: Open the text letters and they are loaded in the template which is compiled.
Step6: Text analysis is done for the combining of word take place.
Step7: After that text to speech library is read those words.
Step8: That finally speech signal is been obtain using MATLAB.

IV. CONCLUSION

This paper reflects the effort being carried out for the text is extracted from the image using optical character recognition and the recognized character is converted to audio using the MATLAB environment. This application proposed is a cost-effective, user compatible, and real-time system. By this system, read a text file from documents, newspapers, E-mail. This system can also be advantageous for people who are visually handicapped. And the more interesting fact is I'm trying to make one hardware space where MATLAB. But the program can be installed and can be carried everywhere easily. The main purpose of this system is to fulfill the needs of the People who are vocally handicapped and also can have communication with people who do not understand the sign language. Another help is one can learn the pronunciation of the words correctly. TTS systems also can be used in domain-specific applications such as train announcements. The system can be used to build information surfing for the people who can't write and read. My project work may help blind people to read the documents and books deaf will get benefited to share their feelings /opinion. The practical experiment is performed text reading system and better results are obtained .at last project seems to be easy but the real-time problems where inevitable like making templates of every character with every font and making it real-time

REFERENCES

[1] Avanti B, Epstein C L, Grossman M, Gee J C “Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain” J. Med. Image Anal.12 1 26-41,2008.
[2] Barb a J L, Vargas Q L, and Torres C et al. 2008 “Micro calcifications detection system through discrete wavelet analysis and image enhancement techniques” Proc. Int. Symptom System Theory (Orlando) p118–121.
[3] Loris Nanni and Alessandra Lumina “A hybrid wavelet-based fingerprint matcher Pattern” Recognition. 40 113146-3151.
[4] Hong L Z, Fang D H, et al. 2010 “The application of digital image recognition to the analysis of two-dimensional fingerprints” Ana. Chem. Act 657 2131-135.