

Raster Format Mapping onto Adobe Reader for Visually Impaired

Gururaj K S

Associate Professor, Department of Information Science and Engineering
GSSS Institute of Engineering & Technology for Women, Mysuru
gururaj.k.s79@gmail.com

Abstract - Texts are scripts provided by the language to convey information. The main idea of this paper is to provide a means for people with vision disabilities to listen to the text of an image by converting text into speech. This process uses text-to-speech, the generation of synthesized speech from text. Image containing the text is pre-processed that prepares the text for recognition. Text recognizer needs to localize the text in image accurately. Separation of text from each other is done by segmentation. Segmentation is followed by the extraction of character, resizing them and storing them in a PDF document. Frequency speech synthesizer provides the speech by performing natural language processing and digital signal processing of the text.

Keywords—Text-recognition, character extraction, segmentation, speech synthesizer

I. INTRODUCTION

Language is the ability to express one’s thoughts by means of a set of signs (text), gestures, and sounds. These texts are read by humans through visualization. People with vision impairment will face difficulty in reading these texts. To amend to the same, speech is generated for the corresponding text. The audio form of the text is known as speech. “Speech synthesis” also called “Text to speech synthesis” is the artificial production of human speech. A computer system used for this purpose is called a speech synthesizer and can be implemented in software. A text-to-speech (TTS) system converts text to speech.

Image acquisition, recognition and speech conversion using Image Write protocol and Text to Speech synthesizer (TTS) by MATLAB is an Image-Processing Technology used to convert the image containing horizontal text into PDF documents and the extracted text is converted into speech.

II. PROBLEM STATEMENT

There are many Text-to-speech converters for the text file and very few to extract the text from image and then generates the speech using OCR.

There is no option to store the extracted text from the image and to restrict the manipulation of the text stored while generating the speech for the same.

Providing the speech of the complete document accurately is still a major challenge for the various existing synthesizers.

This project considers the text image as input and extracts the text using image write protocol. This extracted text is written to a PDF file and generates the speech using frequency speech synthesizer.

III. METHODOLOGY

There are different modules to generate TTS for the given text image mapped on to PDF file. The figure.1 shows the process flow of this project.

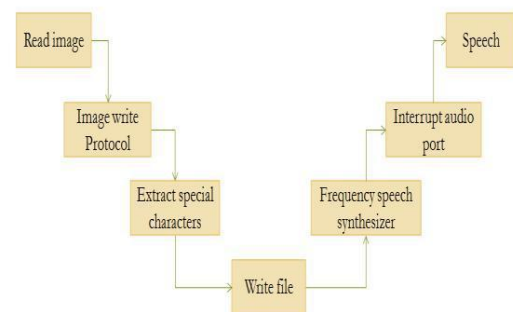


Fig.1: process flow for the TTS

a) Read image

The image that contains printed text is given as input to be read by placing it in the created GUI in MATLAB.

b) Image write protocol

In this phase preprocessing and extraction of characters are done. The aim of preprocessing is an improvement of the image data that suppresses unwanted distortions or enhances some image features that is important for further processing.

c) Special character extraction

Special characters are the symbols other than the alphabet and numbers. These special characters are recognized by the extraction of the features. Feature extraction is one of the important phases. Extracting preliminary features and dividing them into geometric elements and comparing these elements with known set of characters which are stored in the database.

d) PDF file

The texts extracted from image using the MATLAB will be saved as a PDF file. The text from this PDF file is taken as in input for the frequency speech synthesizer.

e) Frequency speech synthesizer

A Text-To-Speech (TTS) synthesizer is a computer based system that should be able to read any text aloud, when it is directly introduced in the computer by an operator. It is more suitable to define Text-To-Speech or speech synthesis as an automatic production of speech by 'grapheme to phoneme' transcription.

A grapheme is the smallest distinguishing unit in a written language. It does not carry meaning by itself. Graphemes include alphabetic letters, numerical digits, punctuation marks, and the individual symbols of any of the world's writing systems. The phoneme' transcripts contain phoneme and prosody.

A phoneme is "the smallest segmental unit of sound employed to form meaningful utterances". Prosody is concerned with those elements of speech that are not individual phonetic segments that are vowels and consonants but are properties of syllables and larger units of speech.

These scripts are processed and generate a wave file for the texts corresponding text depending on the databases created.

Creation of database:

Building the unit inventory consists of three main phases. First, the natural speech must be recorded so that all used units (phonemes) within all possible contexts (allophones) are included. After this, the units must be labeled from spoken speech data, and finally, the most

appropriate units must be chosen. Gathering the samples from natural speech is usually very time-consuming. The implementation of rules to select correct samples for concatenation must also be done very carefully. After feature extraction, system requires assistance of database in order to recognize the characters in the image.

f) Interrupt audio port

The audio part must be enabled or interrupted the default audio port by the MATLAB to analyze the wave file generated by frequency speech synthesizer and frame them.

g) Speech

The speech generated for the corresponding text is produced by the enabled default audio port.

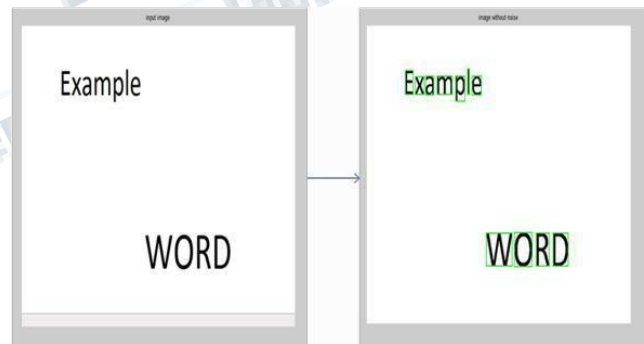
IV.IMPLEMENTATION

a) Imread ()

A = imread(filename) reads the image from the file specified by filename, inferring the format of the file from its contents. If filename is a multi-image file, then imread reads the first image in the file.

b) bwlabel ()

L = bwlabel(BW) returns the label matrix L that contains labels for the 8-connected objects found in BW. The label matrix, L, is the same size as BW.

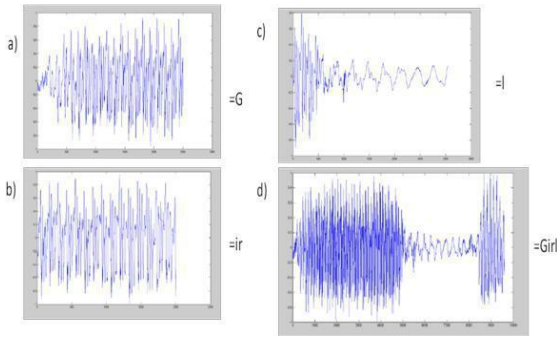


c) Invoke ()

S = invoke(h) returns structure array, S, containing a list of all methods supported by the object or interface, h, along with the prototypes for these methods. If S is empty, either there are no properties or methods in the object, or MATLAB cannot read the object's type library. This function is used invoke getvoices.

d) tts ()

This function is used to convert the text into speech that is stored in the PDF. The text is recognized from the image and given as a parameter to this function which converts it into speech by interrupting the audio port.



V. EXPERIMENTAL RESULTS

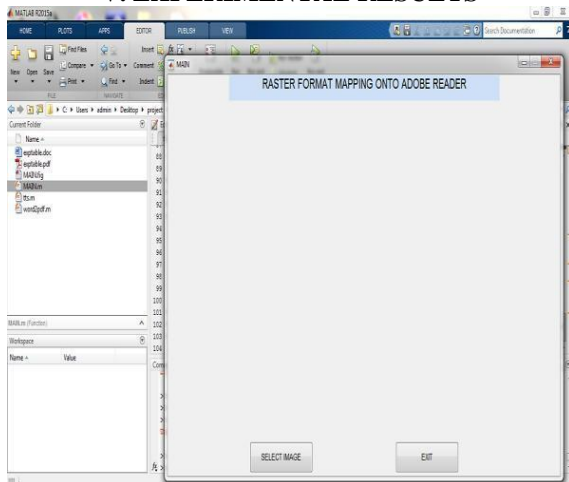


Fig 2.1 shows the Graphical User Interface, through which user interacts with the system by providing the image as input that as to be converted into speech.

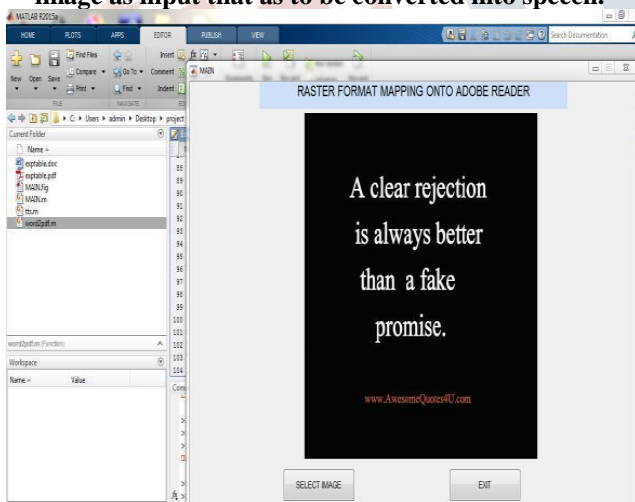


Fig 2.1 shows the snapshot, in which the image stored in the folder is selected by the user is displayed. The texts are extracted from this image and stored in the PDF.

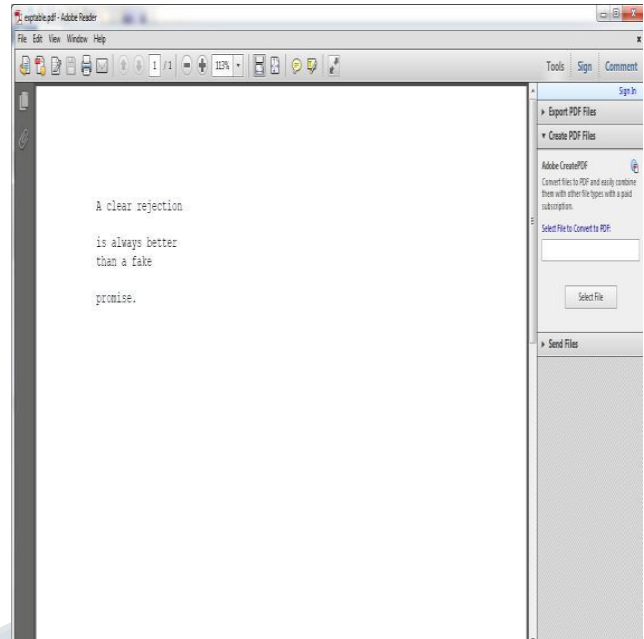


Fig 2.3 shows the snapshot, where the PDF is displayed in which the extracted text from the image is been stored and then its converted into speech.

V. CONCLUSION

Text Image written to PDF and then that text is generated to speech using MATLAB. To recognize the text, the image has to be pre-processed that is, it has to be converted into gray scale image then black and white image from which the characters are separated in segmentation stage. These characters are extracted and written to PDF file by MATLAB. The frequency speech synthesizer generates the speech for given text while comparing with the database created. The default audio port is enabled by MATLAB that allows the speech. The application developed is user friendly, cost effective and applicable in the real time.

REFERENCES

[1] "A Review on Conversionl of Image to Text As Well As Speech Using Edge Detection and Image Segmentation"- Mrunmayee Patil, Ramesh Kagalkar, International Journal of Science and Research (IJSR), ISSN (Online): 2319-7064.

[2] "Text Location in Complex Images" Alvaro Gonz'alez, Luis M. Bergasa, J. Javier Yebes, Sebasti'an Bronte, 21st International Conference on Pattern Recognition (ICPR 2012) November 2012.

[3] "Design and Implementation of Speech

Generation System using MATLAB” Pooja Chandran, Aravind S, Jisha Gopinath and Saranya S S ,International Journal of Engineering and Innovative Technology (IJEIT) Volume 4, Issue 6, December 2014.

[4] “Survey of Methods for Character Recognition Suruchi G. Dedgaonkar, Anjali A. Chandavale, Ashok M. Sapkal, International Journal of Engineering and Innovative Technology (IJEIT) Volume 1, Issue 5, May 2012.

[5] “High Quality Text to Speech Synthesizer using Phonetic Integration” Mrs. S. D. Suryawanshi, Mrs. R. R. Itkarkar, Mr. D. T. Mane, International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE) Volume 3, Issue 2, February 2014

[6]“Real time implementation of image recognition and text to speech conversion” K.Kalaivani, R.Praveena,V.Anjalipriya,R.Srimeena,International Journal of Advanced Engineering Research and Technology (IJAERT) Volume 2 Issue 6, September 2014, ISSN No.: 2348 – 8190

[7]“Implementation of Text to Speech Conversion Technique” Mohd Bilal Ganai, Er jyoti Arora International Journal of Innovative Research in Computer and Communication Engineering (*An ISO 3297: 2007 Certified Organization*) Vol. 3, Issue 9, September 2015

[8]“Text to Speech Conversion System using OCR” Jisha Gopinath¹, Aravind S, Pooja Chandran³, Saranya S S, International Journal of Emerging Technology and Advanced Engineering (ISSN 2250-2459, ISO 9001:2008 Certified Journal Volume 5, Issue 1), January 2015.

[9]“Review of text to speech conversion methods Poonam.S.Shetake, S.A.Patil, P. M Jadhav, International Journal of Industrial Electronics and Electrical Engineering, ISSN: 2347-6982 Volume-2, Issue-8, Aug.-2014