

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 7, Issue 5, May 2018

## Talking Eye for Blind People

Kavita Wani<sup>1</sup>, Sayali Shinde<sup>1</sup>, Prof.S.G.Dube<sup>2</sup>

B. E Student, Department ENTC, MKSSS'S Cummins College of Engineering for Women, Karvenagar, Pune, India<sup>1</sup>

.MKSSS'S Cummins College of Engineering for Women, Karvenagar, Pune, India<sup>2</sup>

**ABSTRACT:** A speech synthesis synthesizer is an application that converts text into spoken words, analyzing and processing text using natural language processing and then using digital signal processing technology to convert this processed text in a synthesized speech representation of the text. Here, we develop a useful text-to-speech synthesizer in the form of a simple application that converts text entered into speech synthesis and reads it to the user, which can then be saved as an mp3 file. The development of a text-to-speech synthesizer will be of great help to people with visual impairments and will facilitate the creation of large volumes of text.

Optical character recognition (OCR) is the identification of printed characters using photoelectric devices and software. Hides text images written, handwritten or printed in machine-encoded text from a scanned document or subtitle text superimposed on an image. In this investigation, these images become audio output. OCR is used in machine processes such as cognitive computation, machine translation, text to speech, key data and text extraction. It is used mainly in the field of research in character recognition, artificial intelligence and artificial vision. In this investigation, since the recognition process is performed through OCR, the character code in the text files is processed using a Raspberry Pi device in which it recognizes characters using tesseract, the programming of the algorithms and pitons and the audio output.

**KEYWORDS:** Text-to-speech synthesis, Natural Language Processing, Digital Signal Processing ,Energy efficient Character recognition , document image analysis(DIA), Raspberry PI, audio output, OCR based book reader, python programming

### I. INTRODUCTIONS

The text to speech system gets the text like input and then an information algorithm that called the text to speech engine analyze the text, pre-process the text and summarize the text speech with some mathematical models. the text to speech engine Generally generates audio data in audio format as output.The text to speech synthesis procedure consists of two main steps. The first is text analysis, where the input text is transcribed into a phonetic representation or other linguistic representation, and the second is the generation of speech waveforms, in which the result is produced by this phonetic information and prosodic. These two phases are often referred to as high and low level synthesis [1]. A simplified version of this procedure is presented in Figure 1 below. The input text can be, for example, data from a word processor, standard ASCII from an e-mail, a mobile text message or text scanned from a newspaper. The string of characters is then pre-processed and analyzed in a phonetic representation which is generally a string of phonemes with additional information for the correct pitch, duration and tension. The sound of the speech is finally generated with the low-level synthesizer from the high-level information. The artificial production of vocal sounds has a long history, with documented mechanical attempts dating back to the eighteenth century.

OCR technology allows the conversion of scanned images of text or symbols printed in text or information that can be understood or modified using a computer program. In our OCR technology system we use the tesseract library. Using the flit library, the data will be converted into audio. The camera acts as the main vision to detect the image of the product or plate label, then the image is processed internally and separates the label from the image through an open CV library and then identifies the product and the name of the identified product is pronounced through the item. Now it identifies that the image of the received label is converted into text using the tesseract library. Once the identified tag name is converted to text, the converted text is displayed in the screen unit connected to the controller. Now the

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 7, Issue 5, May 2018

converted text must be converted to voice to hear the label name as an entry via headphones connected to the audio input port using the Flit library.

## A. Motivations:

In text to speech we have to convert text data into audio file. pre-process the text and summarize the text speech with some mathematical models. the text to speech engine Generally generates audio data in audio format as output. we used OCR to convert image data into textual format . In our OCR technology system we use the TESSERACT library. Using the Flit library, the data will be converted into audio.

## B. Objective and Scope:

In our system we convert image data into text file as well as into audio file. Same way we convert text file data into audio file. In text to speech conversion we used some machine learning method and preprocessing technique. OCR tool is used to extract textual data from image and normalize it into natural language.

Main objective of our system is given as bellow:

- 1) for e-learning program we used text-speech software.
- 2) we have described a prototype system to read printed text and hand held objects for assisting the blind people.
- 3) Text-to-Speech is guarantees the high quality of the training and e-learning courses makes employees and channel partners the best result and fulfills expectations and requirements.

## II. RELATED WORK OR LITERATURE SURVEY

Xiangrong Chen, Alan L. Yuille .This paper gives an algorithm for detecting and reading text in natural images. The algorithm is intended for use by blind and visually impaired subjects walking through city scenes. We first obtain a dataset of city images taken by blind and normally sighted subjects. From this dataset, we manually label and extract the text regions[1].

KwangIn Kim, Keechul Jung, and Jin Hyung Kim. The current paper presents a novel texture-based method for detecting texts in images. A support vector machine (SVM) is used to analyze the textural properties of texts. No external texture feature extraction module is used; rather, the intensities of the raw pixels that make up the textural pattern are fed directly to the SVM, which works well even in high-dimensional spaces[2].

Xilin Chen. we present an approach to automatic detection and recognition of signs from natural scenes, and its application to a sign translation task. The proposed approach embeds multiresolution and multiscale edge detection, adaptive searching, color analysis, and affine rectification in a hierarchical framework for sign detection, with different emphases at each phase to handle the text in different sizes, orientations, color distributions and backgrounds[3].

DimitriosDakopoulos and Nikolaos G. Bourbakis. This paper presents a comparative survey among portable/wearable obstacle detection/ avoidance systems (a subcategory of ETAs) in an effort to inform the research community and users about the capabilities of these systems and about the progress in assistive technology for visually impaired people. The survey is based on various features and performance parameters of the systems that classify them in categories, giving qualitative-quantitative measures[4].

Alan W Black. over the last ten years. Or at least it appears that this is the case. We have moved from diphones to unit selection. However, although we can produce much more natural sounding examples we have also given up an certain amount of control over what can be synthesized. We have reached the stage where playing a few examples to a non-expert can easily convince them that speech synthesis is a solved problem[5].

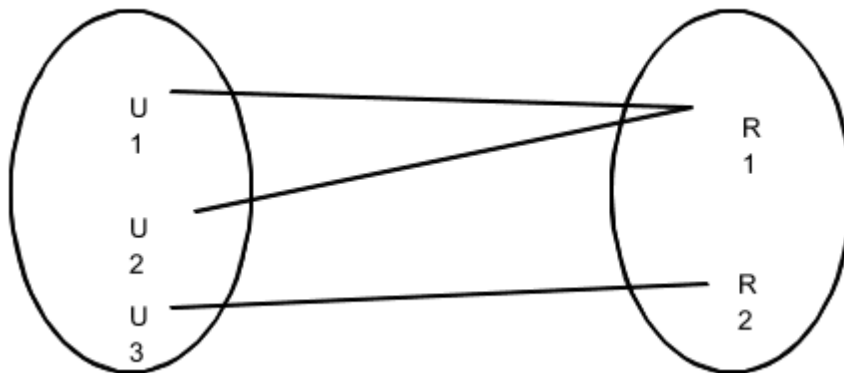
# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 7, Issue 5, May 2018

## III. MATHEMATICAL MODELING



Where,

U=users

R=Result query.

U1=image input file query (R1)

U2= text input file Query (R2)

U3=Wrong input Query (R3)

### B] Set Theory

$S = \{s, e, X, Y, \}$

Where,

s = Start of the program.

1. Log in with web app.

2. select text or image file, convert image or text to audio file. Listen audio file in headset and save audio file in directory.

e = End of the program.

Retrieve all text to speech and image to speech file from directory.

$X = \{V\}$

X = Input of the program.

V = text or audio file.

Y = Output of the program.

First select text or image file, then we will get all text data or image data in array text format. According to user input file system will generate audio file. At final stage user can listen image or text data using audio file reader.

$X, Y \in U$

$U = \{user, V\}$

Where user, V are the elements of the set.

V=text and image file

### Space Complexity:

The space complexity depends on Presentation and visualization of discovered patterns. More the storage of data more is the space complexity.

### Time Complexity:

Check No. of patterns available in the datasets= n

If (n>1) then retrieving of information can be time consuming. So the time complexity of this algorithm is  $O(\quad)$ .

= Failures and Success conditions.

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 7, Issue 5, May 2018

### Failures:

1. Huge database can lead to more time consumption to get the information.
2. Hardware failure.
3. Software failure.

### Success:

1. Search the required information from available in Database.
2. User gets result very fast according to their input file request.

## IV. EXISTING SYSTEM AND DISADVANTAGES

### Disadvantages:

### PROPOSED SYSTEM AND ADVANTAGES

Here We have to identified the various operations and processes involved in text to speech synthesis. Here, we develop a useful text-to-speech synthesizer in the form of a simple application that converts text entered into speech synthesis and reads it to the user, which can then be saved as an mp3 file. The development of a text-to-speech synthesizer will be of great help to people with visual impairments and will facilitate the creation of large volumes of text. Optical character recognition (OCR) is the identification of printed characters using photoelectric devices and software. Hides text images written, handwritten or printed in machine-encoded text from a scanned document or subtitle text superimposed on an image. In this investigation, these images become audio output. OCR is used in machine processes such as cognitive computation, machine translation, text to speech, key data and text extraction. It is used mainly in the field of research in character recognition, artificial intelligence and artificial vision. In this investigation, since the recognition process is performed through OCR, the character code in the text files is processed using a Raspberry Pi device in which it recognizes characters using tesseract, the programming of the algorithms and pitons and the audio output.

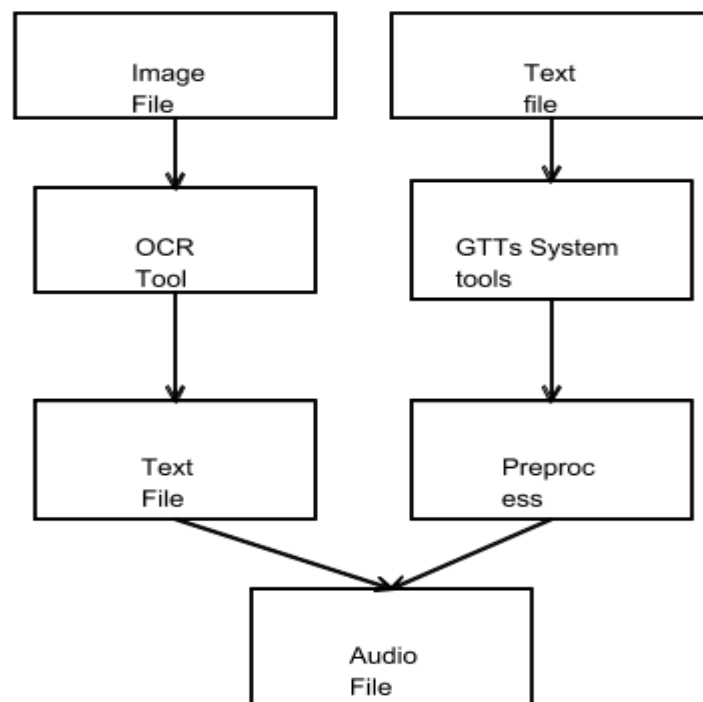


Figure: Proposed System Architecture



ISSN(Online): 2319-8753  
ISSN(Print): 2347-6710

# International Journal of Innovative Research in Science, Engineering and Technology

*(A High Impact Factor, Monthly, Peer Reviewed Journal)*

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 7, Issue 5, May 2018

## V. CONCLUSION

we have described a prototype system to read printed text and hand held objects for assisting the blind people. To extract text regions from complex backgrounds, we have proposed a novel text localization algorithm based on models of stroke orientation and edge distributions. Text to speech synthesis is a rapidly growing aspect of computer technology and is increasingly playing a more important role in the way we interact with the system and interfaces across a variety of platforms. We have identified the various operations and processes involved in text to speech synthesis. InOCR is used to perform word recognition on the localized text regions and transform into audio output for blind users. In this research, the camera acts as input for the paper. As the Raspberry Pi board is powered the camera starts streaming.

## REFERENCES

- 1] Xiangrong Chen, Alan L. Yuille - Detecting and Reading Text in Natural Scenes
- 2] KwangIn Kim, Keechul Jung, and Jin Hyung Kim - Texture-Based Approach for Text Detection in Images Using Support Vector Machines and Continuously Adaptive Mean Shift Algorithm
- 3] Xilin Chen - Automatic Detection and Recognition of Signs From Natural Scenes
- 4] Dimitrios Dakopoulos and Nikolaos G. Bourbakis. Wearable Obstacle Avoidance Electronic travel.
- 5] Alan W Black. Perfect synthesis for all of the people all of the time. IEEE TTS Workshop