IOSR Journal of VLSI and Signal Processing (IOSR-JVSP)
Volume 6, Issue 3, Ver. III (May. - Jun. 2016), PP 15-23
e-ISSN: 2319 – 4200, p-ISSN No. : 2319 – 4197
www.iosrjournals.org

Voice Assisted Text Reading System for Visually Impaired Persons Using TTS Method

Sanjana.B[1], J.RejinaParvin[2]
[1,2]Department of Electronics and Communication Engineering, Dr. NGP Institute of Technology, India

Abstract: *In the survey American Foundation for Blind 2014, is observed that there are 6.8 trillion people are visually impaired and they still find difficult to roll their day today life it is important to take necessary measure with the emerging technologies to help them to live the current world irrespective of their impairments.In the motive of supporting the visually impaired, a method is proposed to develop a self-assisted text to speech module in order to make them read and understand the text in an easier way. It is not only applicable for the visually impaired but also to any normal human beings who are willing to read the text as a speech as quickly as possible.A finger mounted camera is used to capture the text image from the printed text and the captured image is analyzed using optical character recognition (OCR). A predefined dataset is loaded in order to match the observed text with the captured image. Once it is matched the text is synthesized for producing speech output. The main advantage of proposed method is that, it reduces the dataset memory required for the comparison since only character recognition is being done. The same work is simulated using Matrix laboratory (MATLAB) simulator software for the performance analysis of proposed work for various input sets.*

Keywords: *ABF, OCR, Text Reading for Blind, Text to speech method, Visually Impaired Persons*

I. Introduction

Text to speech technology is the process wherein the computer is made to speak. It uses the concepts of natural language processing. In Text reading applications, there are many different techniques available such as label reading, voice stick, brick pi reader and pen aiding but these methods can perform text to speech by creating datasets. In order to address this problem, finger reading technique has been developed, it eliminates the datasets created and stored previously and provide a previous response of reading any text given as input captured image. The speech synthesizer converts the audio input into the text form and processes the text to further learning modules. Despite the advancement of technology that allows for storing information electronically, textual information still remains the most common mode of information exchange. Virtually people who could restore normal vision with eye glasses or contact lenses are around 20% from the survey of ABF (www.abf.com) who could lead their normal lives. Apart from them 90% of world`s visually impaired people who live in low, middle and even in most developed countries, cataract remains the leading cause of blindness.

Accessing text documents is troublesome for visually impaired people in many scenarios, such as reading text on the go and accessing text in less than ideal conditions. The goal is to allow blind users to touch printed text and receive speech output in real-time. The user`s finger is guided along each line via haptic and non-verbal audio cues. The development of such systems requires use of such systems requires use of two technologies that are central to these systems, namely optical character recognition for Text Information Extraction (TIE) and Text-To-Speech (TTS) to convert this text to speech [18]. Text Information Extraction is the first and important function of any assistive reading system and is an integral part of OCR because this process determines the intelligibility of the output speech.

The quality of text-to-speech as well as extending our capabilities to generate expressive synthetic speech. Automatic video text detection and extraction was employed to partition video blocks into text and non-text regions. Recent developments in computer vision, digital cameras, and computers make it possible by developing camera-based products that merge computer vision technology with other existing beneficial products such as optical character recognition systems. OCR is used to recognize words. It can recognize characters, words and sentences without any mistakes. OCR has a high rate of recognition which is the electronic conversion of photographed images of typewritten or printed text into computer-readable text [5] .

Developments in computer technology make it feasible to assist these individuals by developing camera based products. Visually impaired people need some portable assistance to read this printed text. A camera based assistive text reading framework to help blind persons read text labels and product packaging from hand-held objects. From this label and product reading have been extended to read printed text based books for continuous reading. The task performed by device is video processing to extract ROI from video. The video is set to around 5s automatically; boundary detection will be done by comparing the number of pixels [3]. The need to develop a voice assisted text to speech system using optical character recognition method with various input sets and speech output is simulated.

II. Related Works

In Text reading applications there are many different techniques available such as label reading, voice stick, brick pi reader and pen aiding but these methods can perform text to speech by creating datasets as shown in Fig.1. In order to address this problem, finger reading technique has been developed, it eliminates the datasets created and stored previously and provide a previous response of reading any text given as input captured image.

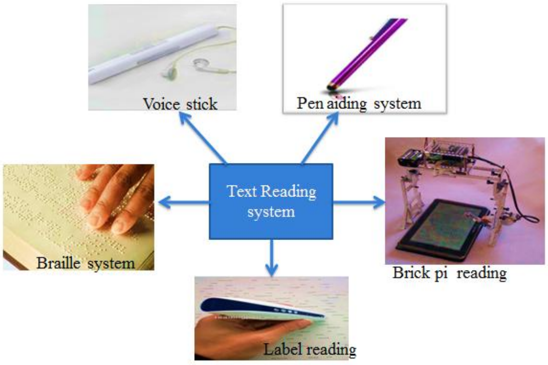


Fig.1 Different text reading Techniques

In paper [1] involving sentence based analysis for expressive of TTS system specifies the problem of creating datasets and by producing output as line by line detection where the VI users find it difficult to recall their sentence. In this scenario, the granularity of text under analysis is usually determined to be the sentence, as sentence are sensibly short textual representations, by regarding this as a sentence by sentence method. Different common features that are used to denote the affect in text are described which matches with the databases stored previously and produces the output aurally. It consist of three modules Input as Text module, Database module and speech module. The input comprises of different features with different affects as lexical fillers, sentence splitter, keyword spotter and word sense disambiguate.

i) Lexical filler converts the plain input text into an output token stream. This module spots the possible affective containers (content words), valence shifters such as negation words and intensifiers.
ii) Sentence splitters splits the sentences in each line which captures the sentence in the form of image and process it aurally.
iii) Word sense Disambiguate resolves the meaning of affective words according to their context. It uses a semantic similarity measure to score the sense of an affective word with the context word using word net ontology.

It uses an ANEW dictionary created already as a set of database. This system can do only by matching with databases created already known as a set of dictionaries. The conditions prescribed above in the input module will be matched with database plotted out in the database section can contain up to 25, 00000 words created in the form of thesaurus. The plain text will be sent to database where it searches for particular word or sentence if the word is correct or found in the dictionary it matches and displays the output if not found it will be sent to trash bin to correct the respective word or sentence.

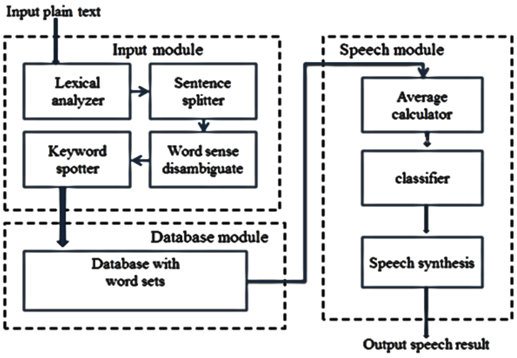


Fig.2 Sentences Based Approach for Text Reading System

A Word set which is matched allows the text to produce the speech synthesis aurally. In speech module, Average calculator, is the arithmetic mean of the dimensions at the sentence level where the matched words will be collected and stored to classify it according to the module it is contained. Classifier predicts the most appropriate sentiment label according to the features extracted from the terms observed in the text, which is usually taken for a bag of words. Finally, the words collected from the bag of words will be sent for speech processing to be left out aurally. This reduces the effectiveness of the classifier to match with necessary datasets required which can hold word set of 25, 0000. The compilation of semeval and twitter dataset constitutes a collection of short sentences with an average of 8 sentences drawn from major newspapers. Formations of typical words have been set already as commonly used words. From the major newspapers, matched words of 250 are processed with a accuracy of 64%. Semeval dataset yields an accuracy of 56% and twitter dataset yields an accuracy of 8%. Additional memories cannot be used to store enough datasets.

The dataset considered is sensibly small with only 250 words matched from the datasets of Semeval and Twitter with the accuracy of 64%. In each dataset there are a set of classifiers and the size of the negative class is more than nine times bigger that the size of the positive class. Each varies with different accuracies.

III. Proposed Voice Assisted Text Reading System

The concept of proposed system is the idea of developing finger reader based text reading system for visually impaired persons. This illustrates the text reading system for visually impaired users for their self-independent.

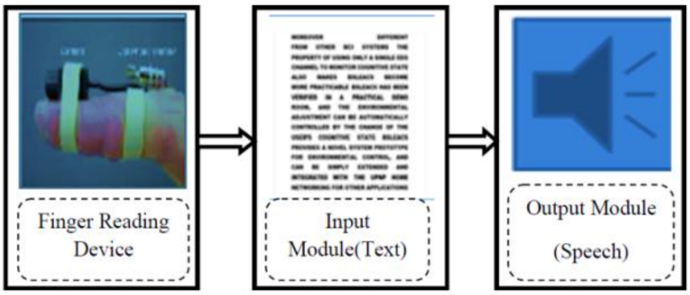


Fig. 3 General Overview of Systems Architecture

The problem stresses the high importance of visually impaired system is that self-dependency of visually impaired users. This extends the work towards the development of ease of collecting information, self-dependent. To achieve the desired result, framework combines a set of different modules, such as finger reading device, TTS module and optical character recognition module. The block diagram of text reading system consist of three parts namely finger reading module, OCR process module& TTS module is shown in Fig The finger reader device is designed with camera, vibration sensor for finger position control.

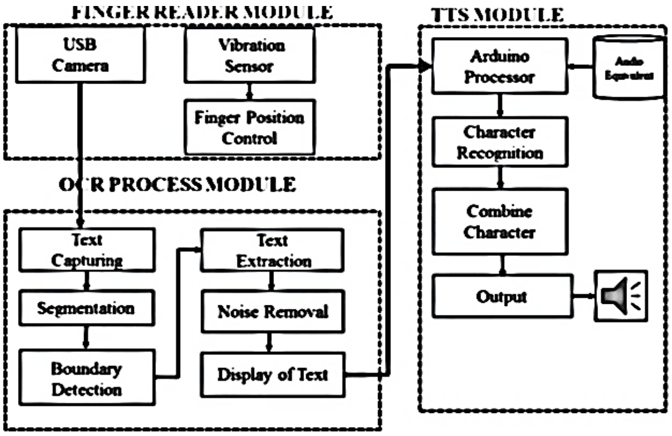


Fig 4 Proposed Voice Assisted Text Reading Block Diagram

A. Finger Reader Module

The reader module was originally developed for reading text formats. The reader module can be a finger-based form factors including small rings. In our current prototype, the camera is attached to an adjustable Velcro ring with the camera inserted on the center of the ring and the vibration sensor which at the corner of the ri g for finger movement and control. For processing, use a wrist-mounted Arduino board with an attached

Bluetooth module that controls the haptic feedback cues. The video feed from the camera is currently processed in real time on a laptop computer.

B. Optical Character Recognition

OCR is optical character recognition module is the mechanical or electronic conversion of images of typed, handwritten or printed text into machine-encoded text. It is a common method of digitizing printed text so that it can be used in machine process such as text-to-speech. OCR is optical character recognition module is the mechanical or electronic conversion of images of typed, handwritten or printed text into machine-encoded text. The input is given as text, using a finger device mounted camera which captures text and sends the input text to the OCR process where the extraction of text to speech is been done. From the captured input text is segmented as word by word detection thereby to read it as separate word. Boundary detection is done by detecting words which are fit inside the boundary, if not it eliminates the text which is unfit to read. The process of text extraction is carried out by matching with templates one by one and then forming a whole word. The mentioned line or a word will be read from the captured input text with a suitable coding. After matching with the templates and displays it as a text and reads it aurally. In this method a USB camera which captures the input given in text format and it is sent to OCR process which processes it as text and converts it into a speech form.

C. TTS Module

A text-to-speech (TTS) system converts normal text into speech other systems render symbolic linguistic representations like phonetic transcription into speech. A text-to-speech system is used to read each word as the user`s finger passes over it, and distinctive audio and/or haptic cues can be used to signal other events, such as end of line, start of line etc. It is composed of two parts: a front-end and a back-end. The front-end has two major tasks. First, it converts raw text containing symbols like numbers and abbreviation into the equivalent of written-out words. This process is often called text normalization, pre-processing, or tokenization the front-end then assigns phonetic transcriptions to each word, and divides and marks the text into prosodic units, like phrases, clauses and sentences. The process of assigning phonetic transcription to words is called text-to phoneme or grapheme-to-phoneme conversion.

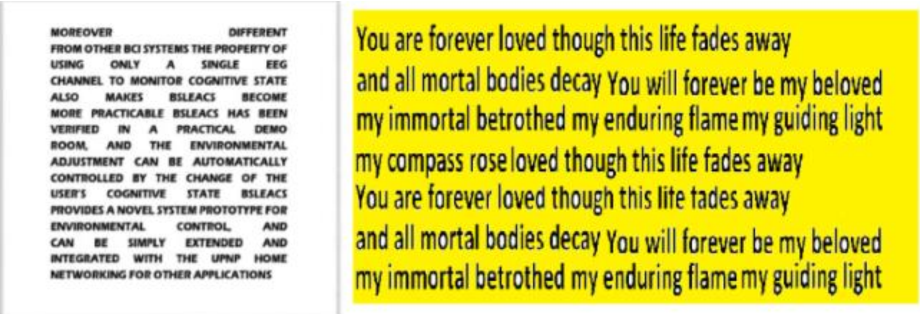
IV. Simulated Results

The proposed system consist of a OCR module, where a USB camera which captures the input given in text format and it is sent to OCR process which processes the text and convert it into a speech form. Captured image is sent to MATLAB, this is to fetch word-by-word segmentation from the input image and compare it with templates. The output of detecting text will be processed from each conversion of RGB TO GRAY, GRAY TO BINARY. The texts in the form of (jpeg, png, jpg, bmp etc) are considered for the analysis. The image which is captured from the USB camera is splited in following conditions as described below for detecting corresponding text and matching it with templates prescribed in the below conditions.

- Text in the form of Black and White
- Text in the form of colored
- Text with image
- Text and image merged
- Text with different Font Styles

B. INPUT IMAGE SETS

The proposed method is analyzed with various text input sets (jpeg, png, jpg, bmp etc). Text in the form of black and white includes a combination of background with plain white and text with black color is shown in Fig (a). The text in the form of colored image contains text with black and colored background as mentioned below in Fig 5.



MOREOVER DIFFERENT FROM OTHER BCI SYSTEMS THE PROPERTY OF USING ONLY A SINGLE EEG CHANNEL TO MONITOR COGNITIVE STATE ALSO MAKES BSLEACS BECOME MORE PRACTICABLE BSLEACS HAS BEEN VERIFIED IN A PRACTICAL DEMO ROOM, AND THE ENVIRONMENTAL ADJUSTMENT CAN BE AUTOMATICALLY CONTROLLED BY THE CHANGE OF THE USER'S COGNITIVE STATE BSLEACS PROVIDES A NOVEL SYSTEM PROTOTYPE FOR ENVIRONMENTAL CONTROL, AND CAN BE SIMPLY EXTENDED AND INTEGRATED WITH THE UPNP HOME NETWORKING FOR OTHER APPLICATIONS

You are forever loved though this life fades away
and all mortal bodies decay You will forever be my beloved
my immortal betrothed my enduring flame my guiding light
my compass rose loved though this life fades away
You are forever loved though this lite tades away
and all mortal bodies decay You will forever be my beloved
my immortal betrothed my enduring flame my guiding light

Fig 5 (a) Black & White image (b) Colored image

The input includes a text with image and the input with text and images merged which is shown in Fig 6 (a) and (b) respectively.

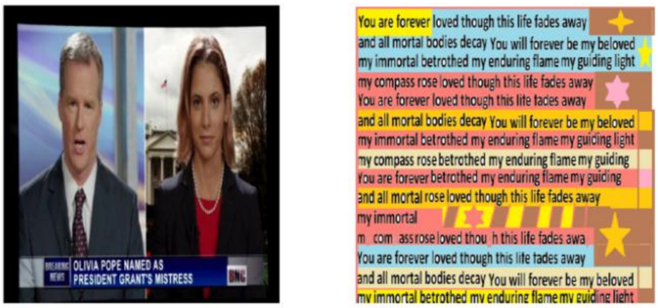
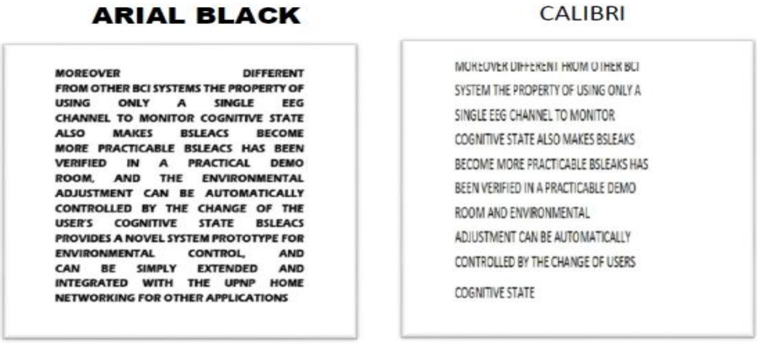

OLIVIA POPE NAMED AS
PRESIDENT GRANT'S MISTRESS
You are forever loved though this life fades away
and all mortal bodies decay You will forever be my beloved
my immortal betrothed my enduring flame my guiding light
my compass rose loved though this life fades away
You are forever loved though this life fades away
and all mortal bodies decay You will forever be my beloved
my immortal betrothed my enduring flame my guiding light
my compass rose betrothed my enduring flame my guiding
You are forever betrothed my enduring flame my guiding
and all mortal rose loved though this life fades away
my immortal
m com assrose loved thou h this life fades awa
You are forever loved though this life fades away
and all mortal bodies decay You will forever be my beloved
my immortal betrothed my enduring flame my guiding light

Fig 6 (a) Text with image (b) Text and images merged

Text with different font styles like Arial black and Calibri are shown in Fig 7 (a) and (b) respectively.



ARIAL BLACK

MOREOVER DIFFERENT FROM OTHER BCI SYSTEMS THE PROPERTY OF USING ONLY A SINGLE EEG CHANNEL TO MONITOR COGNITIVE STATE ALSO MAKES BSLEACS BECOME MORE PRACTICABLE BSLEACS HAS BEEN VERIFIED IN A PRACTICAL DEMO ROOM, AND THE ENVIRONMENTAL ADJUSTMENT CAN BE AUTOMATICALLY CONTROLLED BY THE CHANGE OF THE USER'S COGNITIVE STATE BSLEACS PROVIDES A NOVEL SYSTEM PROTOTYPE FOR ENVIRONMENTAL CONTROL, AND CAN BE SIMPLY EXTENDED AND INTEGRATED WITH THE UPNP HOME NETWORKING FOR OTHER APPLICATIONS

CALIBRI

MOREOVER DIFFERENT FROM OTHER BCI
SYSTEM THE PROPERTY OF USING ONLY A
SINGLE EEG CHANNEL TO MONITOR
COGNITIVE STATE ALSO MAKES BSLEAKS
BECOME MORE PRACTICABLE BSLEAKS HAS
BEEN VERIFIED IN A PRACTICABLE DEMO
ROOM AND ENVIRONMENTAL
ADJUSTMENT CAN BE AUTOMATICALLY
CONTROLLED BY THE CHANGE OF USERS
COGNITIVE STATE

Fig 7 (a) Text with different Font Style (b) Text with different Font Style

The above mentioned input sets are considered and the performance is analyzed using MATLAB R2012a Software. The MATLAB Software is a high-level technical computing language and interactive environment for algorithm development, data analysis and numerical computation. The work flow of text detection and synthesized speech output with different input sets is shown in Fig.8. A particular input set is chosen here among five different input conditions. The condition with black and white text and background is simulated to detect text and to produce a speech output.

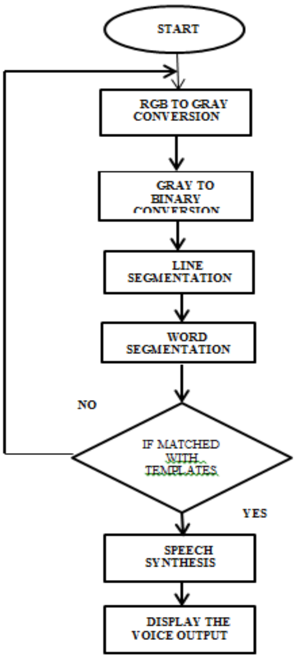


Fig 8 Flowchart for Text detection and Speech synthesizer

Initially the captured input image is fed to software for the analysis the input is in the form of black and white image. The input text image is converted from RGB to GRAY form. In case of colored text image, initially it is converted to Gray form and then to binary in order to segregate 0`s and 1`s, whereas 0`s represents non-text part and 1`s represents the textual part. The converted GRAY image will be then converted into binary form in order to match with the machine. The input image is then segmented in the form of binary. The conversion of binary codes is done by assuming white spaces as 1`s and black spaces as 0`s. The general form of text will be in Matrix form where f1 indicates the first line f2 indicates the second line and so on .The words of the first line can be represented as a_{11}, a_{12}…..a1n where the number '1'indicates the first line and 'n' indicates the number of words in the corresponding line. The general form of representing any word from any line can be given as a_{mn} where, m represents the corresponding line and 'n' represents the corresponding word. Likewise the same form will be followed for representing each line from any part of the paragraph. For example if second word of the third line to read it will be given as $a_{32}$of the third line. The word segmentation is done with the pixel of each letter with the suitable machine language For example, In case, if first word from the first line has to read, it can be segmented as (first line f1 & first word [a_{11}]) where a_{11} represents first word from first line and f1- first line. Similarly, any word from any line can be called for the conversion of text to speech. The segmented word from line will be then matched with a template which is pre-loaded already.

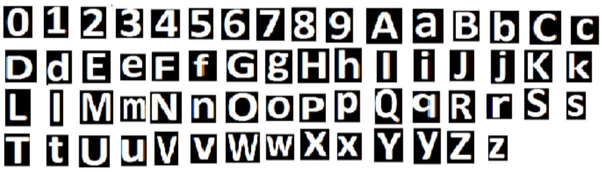
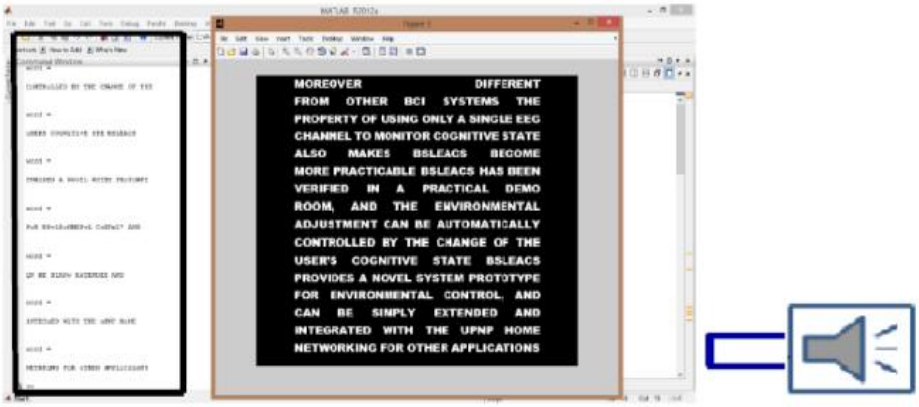

0 1 2 3 4 5 6 7 8 9 A a B b C c
D d E e F f G g H h I i J j K k
L l M m N n O o P p Q q R r S s
T t U u V v W w X x Y y Z z

Fig. 9 Pre-defined Templates

Once the completion of matching with templates, output will be displayed in text format where the input conditions prescribed above will result in text form extracting text alone from the captured input image as shown in below Fig (10) with a box embossed and output will be displays as text in the command window. Through this command window prescribed output to be produced will be displayed as word by word by and aurally intimating the text aside. Thus the procedure for text detection with speech synthesizer is explained in detailed with a workflow of simulated result.



MOREOVER DIFFERENT FROM OTHER BCI SYSTEMS THE PROPERTY OF USING ONLY A SINGLE EEG CHANNEL TO MONITOR COGNITIVE STATE ALSO MAKES BSLEACS BECOME MORE PRACTICABLE BSLEACS HAS BEEN VERIFIED IN A PRACTICAL DEMO ROOM, AND THE ENVIRONMENTAL ADJUSTMENT CAN BE AUTOMATICALLY CONTROLLED BY THE CHANGE OF THE USER'S COGNITIVE STATE BSLEACS PROVIDES A NOVEL SYSTEM PROTOTYPE FOR ENVIRONMENTAL CONTROL, AND CAN BE SIMPLY EXTENDED AND INTEGRATED WITH THE UPNP HOME NETWORKING FOR OTHER APPLICATIONS

Fig. 10 Speech output for Input text

V. Hardware Wireless Interface of Finger Reader

Finger reader module comprises of a camera with a vibration sensor for indication of next line. The input to the camera will be sent to optical character recognition where the process of extraction of text and removal of noise will be taken place. The vocal process (ie) text-to-speech performs the audio equivalent character text from which extracted text is converted into vocal form. The input to the camera is captured as image and sent to arduino processor for MATLAB process which is performed in transmitter side with the help of Bluetooth. At the receiver side, the received image will be again sent to arduino processor for conversion of text to vocal form. In MATLAB process the coding to extract text will be generated and the observed result will be shown in command window at the transmitter side and at the receiver side, received result from the Bluetooth will be shown in arduino COM port with various input sets in printed form are given below

- Text in the form of Black and White
- Text with different Font Styles

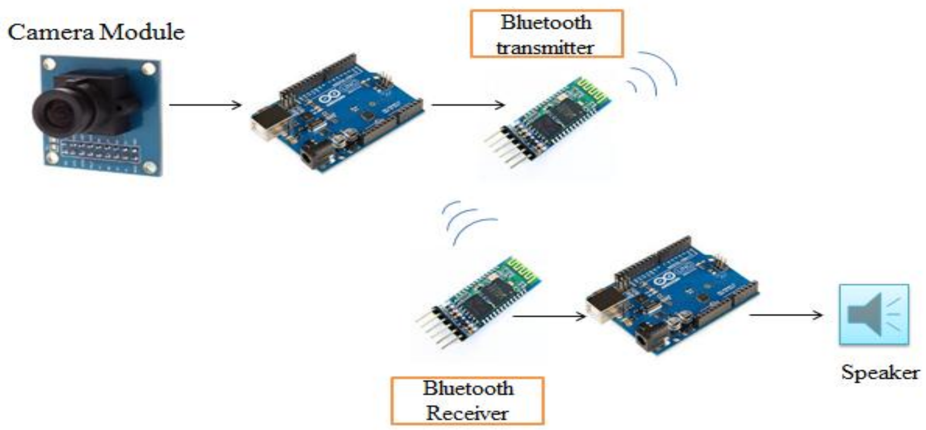


Fig 11 Wireless Interface of Finger Reader

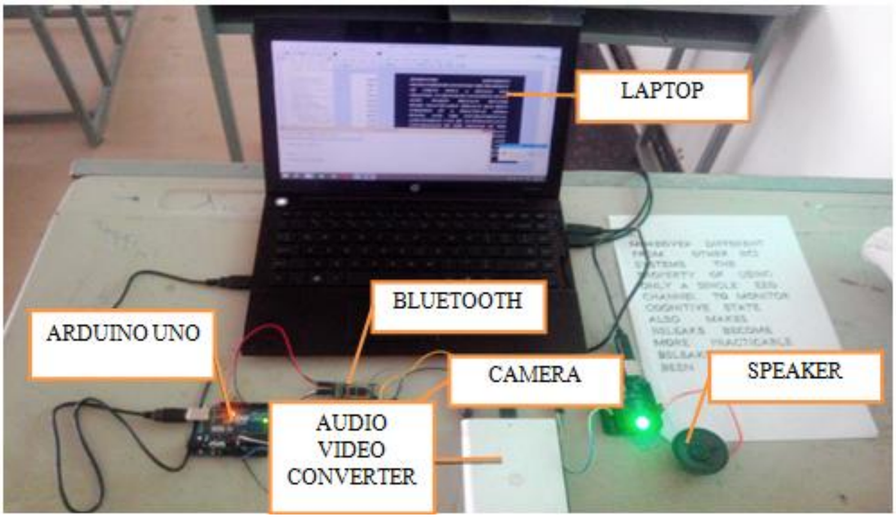
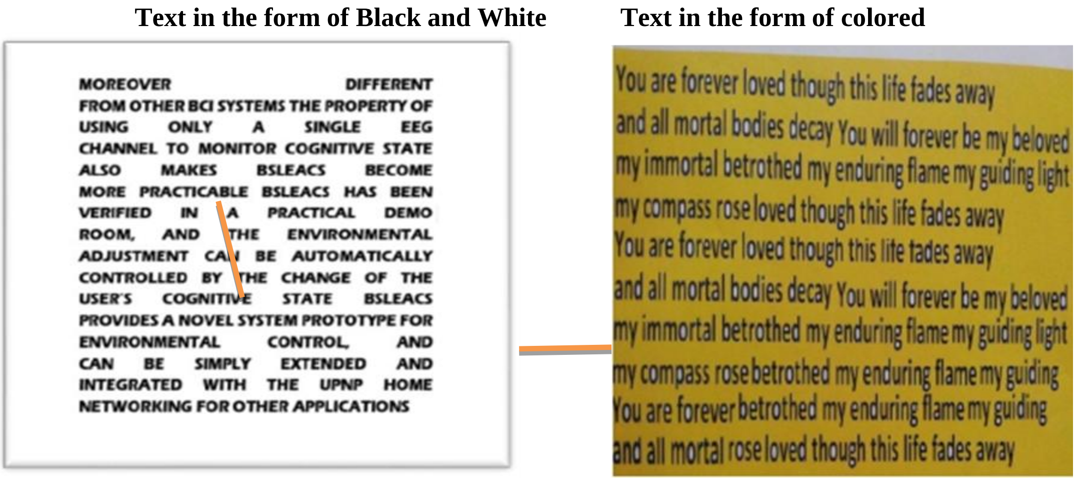


Fig 12 Hardware Snapshot of Text Reading System

Test Results

The proposed system is tested with different input sets as, Text in the form of Black and White, Text with different font styles, Text in the form of colored. In all five sets finger reader is tested using OCR process with around 5 samples of 100 letters which yields an accuracy of around 80%.

Test Results-1



Text in the form of Black and White

MOREOVER DIFFERENT FROM OTHER BCI SYSTEMS THE PROPERTY OF USING ONLY A SINGLE EEG CHANNEL TO MONITOR COGNITIVE STATE ALSO MAKES BSLEACS BECOME MORE PRACTICABLE BSLEACS HAS BEEN VERIFIED IN A PRACTICAL DEMO ROOM, AND THE ENVIRONMENTAL ADJUSTMENT CAN BE AUTOMATICALLY CONTROLLED BY THE CHANGE OF THE USER'S COGNITIVE STATE BSLEACS PROVIDES A NOVEL SYSTEM PROTOTYPE FOR ENVIRONMENTAL CONTROL, AND CAN BE SIMPLY EXTENDED AND INTEGRATED WITH THE UPNP HOME NETWORKING FOR OTHER APPLICATIONS

Text in the form of colored

You are forever loved though this life fades away
and all mortal bodies decay You will forever be my beloved
my immortal betrothed my enduring flame my guiding light
my compass rose loved though this life fades away
You are forever loved though this life fades away
and all mortal bodies decay You will forever be my beloved
my immortal betrothed my enduring flame my guiding light
my compass rose betrothed my enduring flame my guiding
You are forever betrothed my enduring flame my guiding
and all mortal rose loved though this life fades away

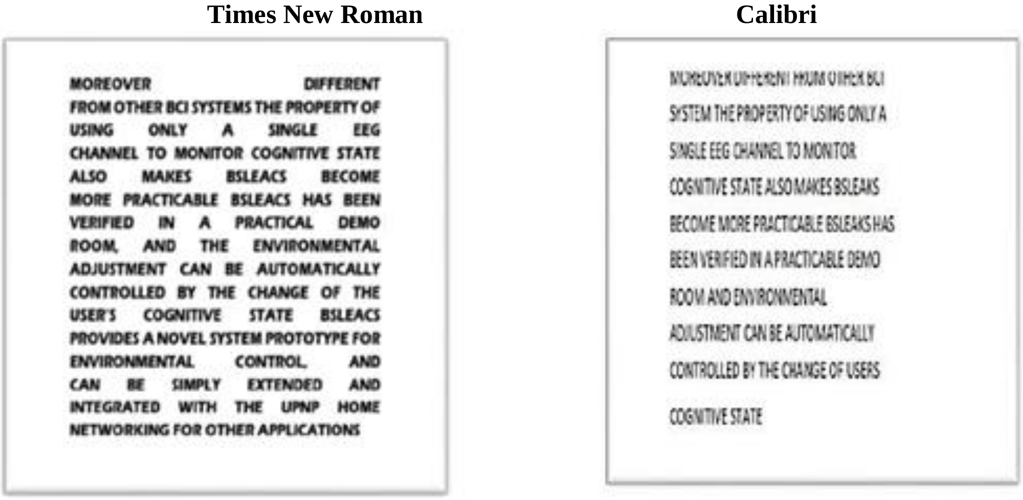
(a) (b)

Fig 13 (a) Black and White (b) colored image

Table 1 Test Results for Black and White and Colored Image

Test results for Black and White and colored image	
No. of Letters	100
Words detected	80
Error possibility	20

Text with Different Font Styles



Times New Roman

MOREOVER DIFFERENT
FROM OTHER BCI SYSTEMS THE PROPERTY OF
USING ONLY A SINGLE EEG
CHANNEL TO MONITOR COGNITIVE STATE
ALSO MAKES BSLEACS BECOME
MORE PRACTICABLE BSLEACS HAS BEEN
VERIFIED IN A PRACTICAL DEMO
ROOM, AND THE ENVIRONMENTAL
ADJUSTMENT CAN BE AUTOMATICALLY
CONTROLLED BY THE CHANGE OF THE
USER'S COGNITIVE STATE BSLEACS
PROVIDES A NOVEL SYSTEM PROTOTYPE FOR
ENVIRONMENTAL CONTROL, AND
CAN BE SIMPLY EXTENDED AND
INTEGRATED WITH THE UPNP HOME
NETWORKING FOR OTHER APPLICATIONS

Calibri

MOREOVER DIFFERENT FROM OTHER BCI
SYSTEM THE PROPERTY OF USING ONLY A
SINGLE EEG CHANNEL TO MONITOR
COGNITIVE STATE ALSO MAKES BSLEAKS
BECOME MORE PRACTICABLE BSLEAKS HAS
BEEN VERIFIED IN A PRACTICABLE DEMO
ROOM AND ENVIRONMENTAL
ADJUSTMENT CAN BE AUTOMATICALLY
CONTROLLED BY THE CHANGE OF USERS
COGNITIVE STATE

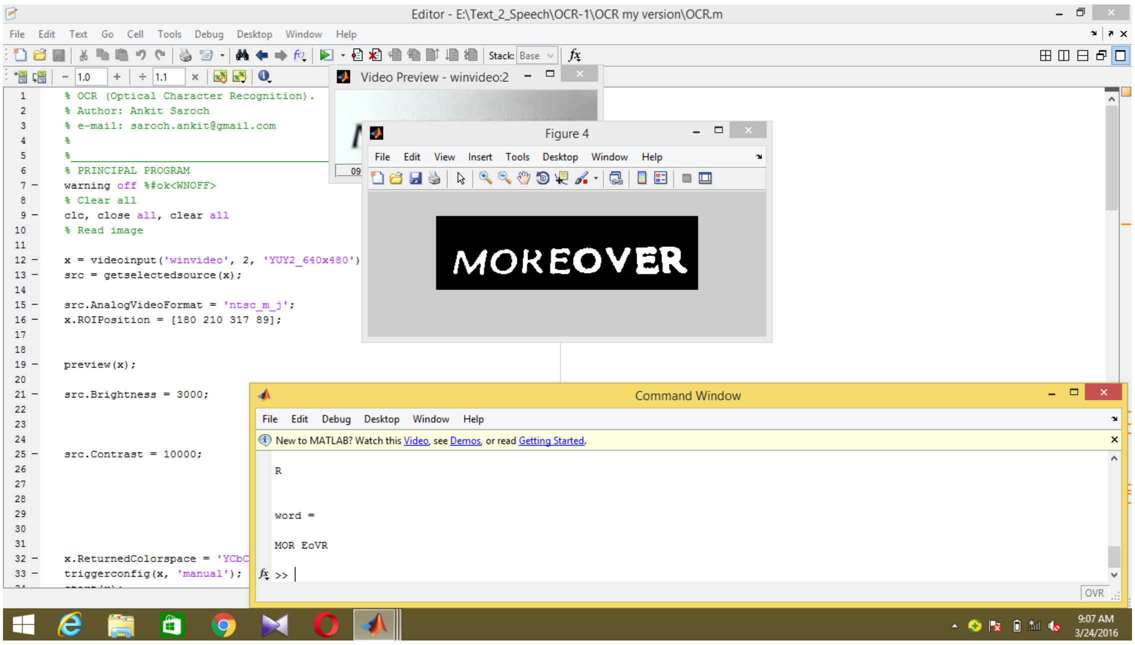
(a) (b)

Fig 14 (a) & (b) Text with Different Font Styles

Table 2 Test results for Text with Different Font Styles

Test results for Text with Different Font Styles	
No. of Letters	100
Words detected	83
Error possibility	17

The word detected result is tested with set-1 Text in the form of Black and White, among different set of words used the first line of first word named a11 is detected shown in fig 15.



Editor - E:\Text_2_Speech\OCR-1\OCR my version\OCR.m
File Edit Text Go Cell Tools Debug Desktop Window Help
Stack: Base
Video Preview - winvideo:2
Figure 4
File Edit View Insert Tools Desktop Window Help
MOREOVER
Command Window
File Edit Debug Desktop Window Help
New to MATLAB? Watch this Video, see Demos, or read Getting Started.
OVR
9:07 AM
3/24/2016

```
% OCR (Optical Character Recognition).
% Author: Ankit Saroch
% e-mail: saroch.ankit@gmail.com
%
%
% PRINCIPAL PROGRAM
warning off %#ok<WNOFF>
% Clear all
clc, close all, clear all
% Read image

x = videoinput('winvideo', 2, 'YUY2_640x480')
src = getselectedsource(x);

src.AnalogVideoFormat = 'ntsc_m_j';
x.ROIPosition = [180 210 317 89];


preview(x);

src.Brightness = 3000;



src.Contrast = 10000;






x.ReturnedColorspace = 'YCbC
triggerconfig(x, 'manual');
```

```
R


word =

MOR EoVR

>>
```

Fig 15 Word Detected Result

Thus the simulated results of various input sets and the hardware setup of finger reader with different input sets and the recognized output is viewed through Matlab simulator with an audio output.

VI. Conclusion

The voice assisted text reading system for visually impaired is discussed. The output is shown for the various input data set like only text inputs, text with images merged etc. Optical Character Recognition is used to predict the input text with pre-loaded database template. Both the characters are compared if it matches then using text to speech synthesizer, speech output is produced. The work is simulated using MATLAB software and the speech output is produced. For the implementation using hardware, a finger module with camera mounted with a vibrator sensor is used. The camera is used to capture the text which is used as an input and the vibration sensor helps to indicate the effective line by line reading. With the help of the proposed module, the user feels easier to read the text in the form of speech using Optical Character Recognition and speech synthesizing.The proposed work is tested with different input sets in printed text format where noise parts are removed and text is extracted to predict the text aurally and is further enhanced by identifying the adjacent character recognition for effective reading to avoid the discontinuity. A output produced as audio output to read the corresponding input which helps the blind people to read any printed text in vocal form.

References

[1]. [AlexandreTrilla and Francesc Alías. (2013),'Sentence-Based Sentiment Analysis for Expressive Text-to-Speech', IEEE Transactions onAudio, Speech, and Language Processing, Vol. 21, Issue. 2. pp. 223-233.

[2]. Alías F. Sevillano X. Socoró J. C Gonzalvo X. (2008), 'Towards high-quality next-generation text-to-speech synthesi', IEEE Trans. Audio, Speech, Language Process, Vol. 16, No. 7. pp. 1340-1354.

[3]. Balakrishnan G. Sainarayanan G. Nagarajan R. and Yaacob S. (2007) 'Wearable real-time stereo vision for the visually impaired', Vol. 14, No. 2, pp. 6–14.

[4]. Chucai Yi. YingLiTian.AriesArditi. (2014), 'Portable Camera-based Assistive Text and Product Label Reading from Hand-held Objects for Blind Persons', IEEE/ASME Transactions on Mechatronics, Vol. 3, No. 2, pp. 1-10.

[5]. Deepa Jose V. and Sharan R. (2014), 'A Novel Model for Speech to Text Conversion', International Refereed Journal of Engineering and Science (IRJES) Vol.3, Issue.1, pp. 39-41.

[6]. Goldreich D. and Kanics I. M. (2003), 'Tactile Acuity is Enhanced in Blindness', International Journal of Research And Science, Vol. 23, No. 8,pp. 3439–3445.

[7]. Joao Guerreiro and Daniel Gonçalves (2014), 'Text-to-Speech: Evaluating the Perception of Concurrent Speech by Blind People', International journal of computer technology, Vol. 6, No. 8, pp. 1-8.

[8]. J. Liang D. and DoermannH. (2005), 'Camera-based analysis of text and documents: a survey,' International Journal on Document Analysis and Recognition, Vol.7, No-6, pp. 83-200.

[9]. Manduchi R. and Miesenberger K. (2012), 'Mobile Vision as Assistive Technology for the Blind: An Experimental Study', Springer-In Computers Helping People with Special Needs, Vol. 2, No.7383, pp. 9–16.

[10]. Marion A. Hersh Michael A. Johnson (2013), 'Assistive Technology for Visually Impaired and Blind', Springer-International Journal of Engineering and Technology, Vol. 4, No. 6, pp. 50-69.

[11]. S. Mascaro. And H. H. Asada. (2001) "Finger posture and shear force measurement: Initial experimentation," in Proc.IEEE Int. conf.robot.Autom, Vol. 2,pp. 1857-1862.

[12]. Norman J. F. and Bartholomew A. N. (2011), 'Blindness Enhances Tactile Acuity and Haptic 3-D Shape Discrimination', Proceedings of the 4thAugmented Human International Conference, Vol. 73, No. 7, pp. 23–30.

[13]. Pitrelli J. and Bakis R.(2006), 'The IBM expressive text-to-speech synthesis system for American English', IEEE Trans. Audio, Speech, Lang. Process, Vol. 14, No. 4, pp. 1099–1108.

[14]. PriyankaBacche. ApurvaBakshi, KarishmaGhiya. PriyankaGujar. (2014), 'Tech –NETRA (New Eyes to Read Artifact)', International Journal of Science, Engineering and Technology Research (IJSETR), Vol. 3, Issue. 3, pp. 482-485.

[15]. Rissanen, M. J.Fernando S.Pang .N. (2013), 'Natural and Socially Acceptable Interaction Techniques for Ringterfaces: Finger-ring Shaped User Interfaces', Springer - In Distributed Ambient and Pervasive Interactions, Vol. 19, No. 6, pp. 52-61.

[16]. RohitRanchal .YirenGuo . Keith Bain and Paul Robinson J (2013), 'Using Speech Recognition for Real-Time Captioning and Lecture Transcription in the Classroom', IEEE Transactions On Learning Technologies, Vol. 6, No.4, pp. 12-17.

[17]. Rubesh Kumar and Purnima (2014), 'Assistive System for Product Label Detection with Voice Output for Blind Users', International Journal of Research in Engineering & Advanced Technology, Vol. 1, Issue. 6, pp. 30-45.

[18]. [Rupali and Dharmale (2015), 'Text Detection and Recognition with Speech Output for Visually Challenged Person', Research Gate- International Journal of Engineering Research and Applications, Vol. 5, Issue. 3, pp.84-87.

[19]. Shen, H. and Coughlan, J. M. (2012), 'Towards a real-time system for Finding and Reading Signs for Visually Impaired Users', Springer-International Journal in Computers Helping People with Special Need, Vol.2 , No. 1, pp. 41-47.

[20]. Shinnosuke and Takamichi (2014), 'Parameter Generation Methods With Rich Context Models for High-Quality and Flexible Text-To-Speech Synthesis' , IEEE Journal Of Selected Topics In Signal Processing, Vol. 8, Issue. 2, pp. 239-250 .

[21]. Tapas Kumar Patra and Biplab (2014) ,'Text to Speech Conversion with Phonematic Concatenation', International Journal of Electronics Communication and Computer Technology (IJECCT) Vol. 2, Issue. 5. pp.223-226.

[22]. Xiang Peng. Fang Liu.Tianjiang Wang. Songfeng Lu (2008), 'A Density-base Approach for Text Extraction in Images', Research Gate- International Journal of Communication and Engineering Vol. 8, No. 4, pp. 239-248.