International Journal of Emerging Trends & Technology in Computer Science (IJETTCS)
Web Site: www.ijettcs.org Email: editor@ijettcs.org
Volume 6, Issue 3, May- June 2017 ISSN 2278-6856

Artificial Vision for Blind Peoples Using OCR Technology

Ramkishor V [1], Mr.Rajesh L [2]

[1] M.Tech, CNE, Dept. of ISE, SJBIT, Bengaluru,

[2] Asst.Professor, Dept. of ISE, SJBIT, Bengaluru,

Abstract: ***Among the total population of world, nearly 270 million people are blind and 48 million people have no proper vision. The disability to read text visually has its effect on personal life of blind people through several instruments are develop to enable the blind people to visual things through sense of sound and touch, The design of devices that reads text is still in the initial stage. The present system for reading text has number of limitation. These systems recognize only few particular shapes or colors. Sometimes they required user assistance and they are expensive too. Hence it is necessary to develop a system that has built a technology to identify and read text a loud to blind people. The theme of the project is to develop a device that can identify text of Kannada, Tamil and English languages change it into voice and audio. The textual characters are Pre-processed by a module which also recognized it. The letters in the text are separated followed by the resizing and story in text file. All these are achieved with the assistance of MATLAB. Finally, the recognized text is heard in the form of speech by the visually impaired people.***

Keywords: Algorithm; Optical Character Recognition (OCR); Segmentation; Text to Speech (TTS)

1. INTRODUCTION

The 3,88 percent of total population that is 270 million people of the world are blind. Till the last decade, braille helped the visually impaired people to identify text. Major backlog of Braille is that only few volumes are modified into braille and only few people can read braille. This application helps blind people to read printed text has normal person read. This project presents on TTS system designed especially for Kannada, Tamil and English languages. Text-to-speech (TTS) convention transforms linguistic information stored as data or text into speech. It is widely used in audio reading devices for blind people now a day's [1]. Also people who experience dyslexia, reading challenges, or visual impairment. There may be an obvious need that leads to the introduction of TTS into an individual's lifestyle, reducing eye strain from reading (digital or paper formats), Reducing paper use due to printing digital text, foreign language learning, writing and editing, or promoting listening skills. A speech synthesizer can be implemented by both hardware and software.

According to the speech generation model used, speech synthesis can be classified into three categories as Articulatory synthesis, Formant synthesis and concatenate synthesis [3]. Concatenate synthesis simply plays back the waveform with the matching phone string. An utterance is synthesized by concatenating together several speech fragments, unlike synthesis-by rule [2]; it requires neither rules nor manual tuning. Moreover, each segment is completely natural, so we should expect very natural output. Speech segments are greatly affected by co-articulation, so if we concatenate two speech segments that were not adjacent to each other, there can be spectral or prosodic discontinuities.

2. PROBLEM STATEMENT

The difficulties encountered while developing this application are as follows:

a. Build an algorithm that can automatically read a text.
b. CPU, Camera resolution and RAM limitations.
c. Unnecessary noise due to reflection or shadows.
d. Variation in the font size, font style and languages.

3. SYSTEM DESCRIPTION

The constituents of the system are:

1) Image processing unit
2) Pair of glasses equipped with HD camera
3) Bluetooth Headset

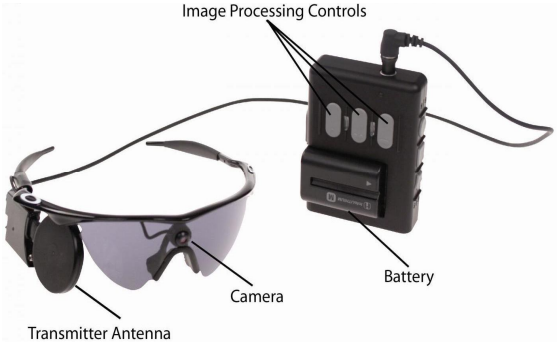


Figure 1 Typical use of Proposed System

4. PRIOR WORK

Many attempts have been made to convert black and white images to voice format. This includes octophone and optocon Optical Character Recognition and Text to Speech technologies are used to develop certain application and devices that help blind people to read text [1].

5. ALGORITHM

The purpose is to develop a device that gives natural experience to the user. A pair of glasses is used .A micro camera and a processing unit is hidden back can be held normally and need not be more curved. Based on this criterion, an algorithm was developed.

It has four phases:

1) Pre –processing
2) Segmentation
3) Feature extraction
4) Recognition

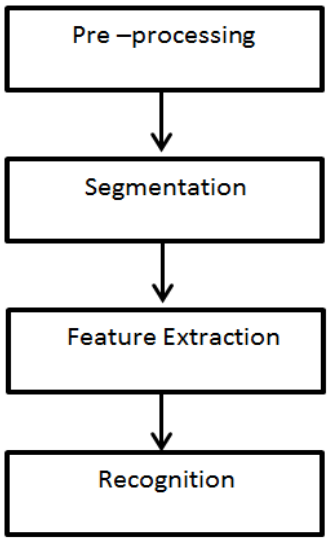


Figure 2 System Flow

5.1 Pre - Processing

The Process of image in a mathematical operation using an Image Processing Unit which consists an image as an input. Process involves isolating the individual color planes of an image and treating them as two-dimensional signal. The aim of pre-processing is an improvement of the image data that suppresses unwanted distortions or enhances some image features important for further processing.

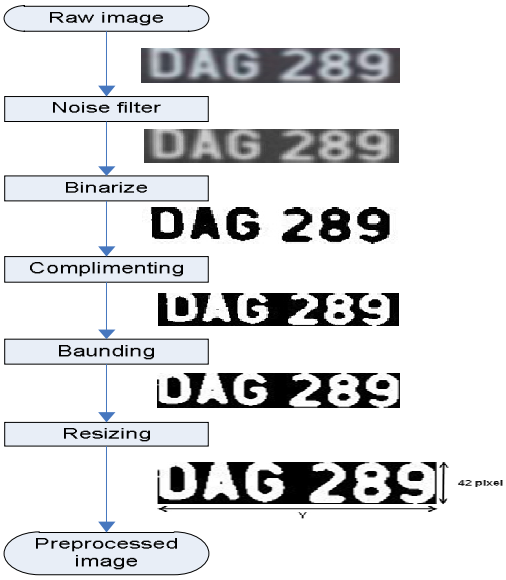


Figure 3 Flow of Pre-Processing

5.2 Segmentation Process

This process mainly consists of line segmentation, word segmentation and then the characters. These segments are then put together for proper identification of the text [2]. Individual letters in a world are separated in Kannada and Tamil languages as they are non-cursive script.

5.2.1 Line Segmentation

Recognition of lines in the given image has following steps:

- Text image is scanned horizontally along the vertical co-ordinate axis to find the first ON pixels.
- We find lots of ON pixels on scanning the text image because the characters would have started.
- Along the horizontal co-ordinate axis we find the first OFF pixels.
- Line formed is from vertical to horizontal.
- Final image is obtained by repeating the above steps.

5.2.2 Word segmentation

Scanning the image vertical, we can segment the line into words .it involves the following steps:

- To find the first ON pixel, the line segment of the text image is scanned vertically along x_1 axis.
- Once the word would have started, we find number of on pixels.
- The final OFF pixels column is obtained along the x_2 co-ordinate axis.
- Word formed is from x_1 to x_2.
- The steps are separated to obtain the complete line segment.

5.2.3 Character Segmentation

- The required word segment is scanned vertically along x_1 co-ordinate axis to get first on pixel.
- . Once the character start, we find lots of on pixel.
- Along the x_2 co-ordinate axis we obtain the final off pixel column.
- Character is from x_1 to x_2.
- Above steps are separated till the end of line segment.

5.3 Text Synthesis

By the use of optical character recognition scripts can be easily identified. The components of this system include scanner, binarization device, segmentation, extraction and identification. The text is obtained through a scanner and is segmented. The scanned features are recognized and the words are reconstructed.

5.3.1 Text to speech synthesis

This is a system that will read the detected text to the user. The block diagram below explains this synthesizer.

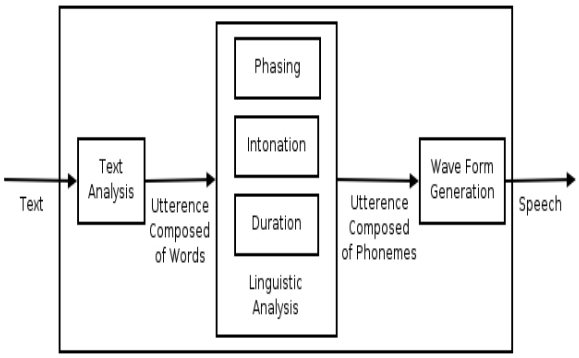


Figure 4 Text to speech synthesis

The system is consists a two parts of front end and Back end. The front end converts symbols, numbers into equalent of written words. This is called pro-process or normalization. The back end is converts symbols or text in to voice.

6. System implementation using MATLAB

The MATLAB is a high level matrix language. It controls the functions flow statement, data structure, input/output. It consists the “Programming in the small” to quick and throw away programs, and “Programming in the large” to create complex application programs [1]. The MATLAB image processing toolbox for this software, it involves changes the nature of image in order to text by a processing toolbox. The process toolbox supports large operation.

Image processing for the blind people is done using the OCR engine tool. Required image is selected from device or a new image is captured. This is rectified and the input is passed to tesseract.

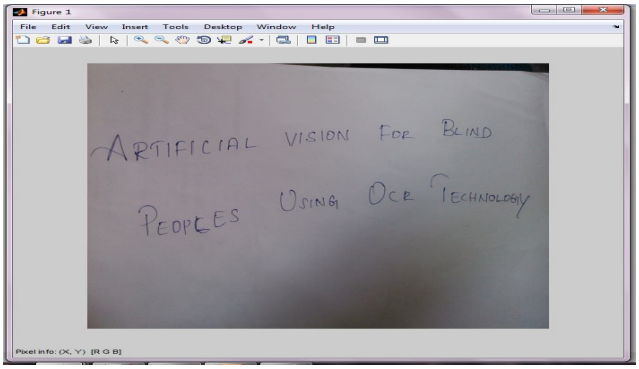
The image processed is displayed and the user can also edit it using TTS. The system can converts as audio of English, Kannada, and Tamil languages. The required languages is selected for proper voice and dictionary. The system consists of an image processing device, a camera for capturing and a headset. Image processing system is used for speech synthesis. The capacity of the developed system is 100 MHz and 16 GB RAM [5]. These are the minimum requirement for this software. A speaker is used for heard the sound.

7. RESULT

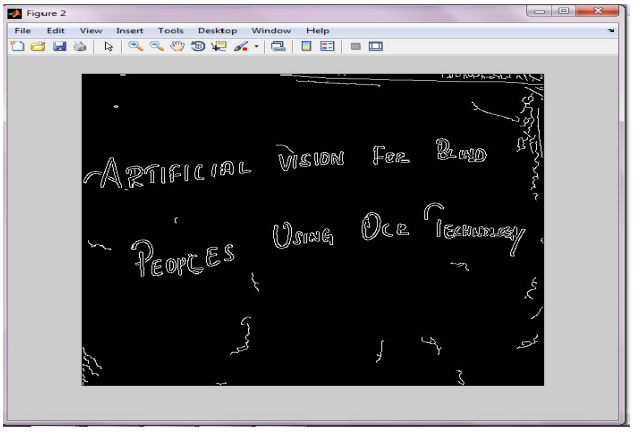
Conversion of image to text and text to voice are the two parts of this text reading system and this is done using MAT LAB. This application converts image into gray image and then into black and white image .speech information is produced by speech application program interface by opting the voice and audio of our choice. We can listen to the opted voices from list using headset.

Using Mat Lab

Input:

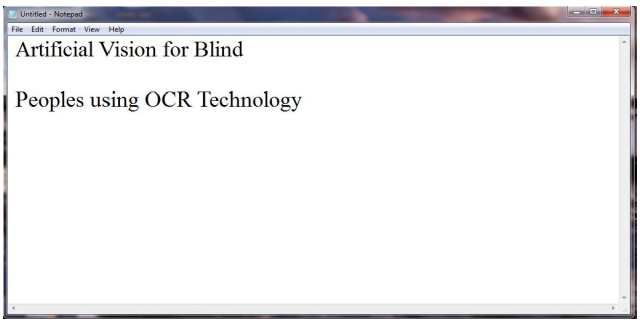


Input image



Text Reading System

Output:



Output Text

8. Applications

Text to speech should be made audible communicate information to the user, when digital audio recordings are inadequate, for developing a user friendly speech synthesizer. Thus this system widely helps in developing a Computer Human interaction like-voice annotations to files, Speech enabled applications, talking computer systems (GPS, Phone based) etc.

Nowadays Text To Speech (TTS) technology is advancing at a faster rate in making people multitasking. TTS on a computer refers to the combination of text appearing on the computer display together with the computer speaking that text aloud with a digitized or synthesized voice. The paper describes a method that uses MATLAB for the recognition of wave files (.wav) of phonemes and array of different phonemes is created. The input text is disintegrated into letters and corresponding phonic sounds are traced using array created and are concatenated using simple matrix concatenations and the desired speech output is fed to the system speaker.

9. Conclusion

The complete effort of this paper is the suggestion to develop a technology that can convert picture into textual characters and then into speech signal, this application can be used by VI peoples. This technology can be used to identify text from any source and speech is prepared through a speaker of phone or computer. The methodology of pronouncing alphabets is developed in the device along with proper grammar and dictionary of Kannada, Tamil and English languages. The system is also provided with an additional advantage of browsing information which is of real use for people with disability to read. There is possibility to convert information both text and picture. This device is of great use to people with visual disability to read textual characters. Not only people with blindness, but also people who cannot speak or dumb can make use of this approach to convert typed words into voice. Positive results have been obtained for the experiments conducted to test this device.

Acknowledgement

I Extend My Soulful Thanks and Gratitude to the Academic Staff of Information Science and Engineering Department of Visvesvaraya Technological University For all their teaching, Guidance and Support provided throughout the Course.

References

[1] Christos Liambas, Miltiadis Saratzidis “Autonomous OCR Dictating System for Blind People” , 978-1-5090-2432-2/16/$31.00 ©2016 IEEE.

[2] B.M.SAGASR, Dr. SHOBHA G, Dr RAMAKANTH KUMAR P, "OCR for printed Kannada text to Machine editable format using database approach. VTU 24-26,2008

[3] Jisha Gopinath, Aravind S, Pooja Chandran, Saranya S S , "Text to Speech Conversion System using OCR" SBCEW ,kerala , india, volume 5, issue 1, january 2015,

[4] DJ.Ravil and Sudarshan Patilkulkarni " Kannada Text to speech Synthesis Systems : Emotion Analysis" in the Proceedings of the seventh International Conference on Natural Language Processing (rCON 2012), Hyderabad, held on 14th to 17th Dec 2012, pp 51-58.

[5] M. Clerc, “The Swarm and the Queen: Towards a Deterministic and Adaptive Particle Swarm Optimization,” In Proceedings of the IEEE Congress on Evolutionary Computation (CEC), pp. 1951-1957, 1999. (conference style)

[6] X. Huang, A.Acero and H.-W. Hon, "Spoken Language Processing A Guide to Theory, Algorithm and System Development", New Jersey : Prentice Hall, 2015.

[7] H.H. Crokell, “Specialization and International Competitiveness,” in Managing the Multinational Subsidiary, H. Etemad and L. S, Sulude (eds.), Croom-Helm, London, 1986. (book chapter style).