

# Smart Reading Assistance

Purushottam Musale  
Extc dept.  
ACE  
Mumbai, India

Deepak Hegde  
Extc dept.  
ACE  
Mumbai, India

Riya Kolge,  
Extc dept.  
ACE  
Mumbai, India

Pratik Pawar,  
Extc dept.  
ACE  
Mumbai, India

Professor Jyoti kolap  
Extc dept.  
ACE  
Mumbai, India

**Abstract**— A Majority of the people in India are visually impaired and blind. This gives rise to the need for the development of devices that could bring relief to them. This paper aims to study the technology of image recognition with speech synthesis and to develop a cost effective, user friendly image to speech conversion system with help of Matlab. The paper includes system which has a inbuilt small camera that scans the text printed on a paper, converts it into audio format using a synthesized voice for reading out the scanned text quickly translating books, documents and other materials for daily living, especially away from home or office (TTS). Finger tracking based a virtual mouse application has been designed and implemented using a regular webcam. Not only it saves time and energy, but also makes life better for the visually impaired as it increases their independency

**Keywords**— *Recogniton, Finger tracking, Speech synthesis, TTS*

## I. INTRODUCTION

Over the last few decades, machine reading has grown from a dream to reality. Speech is probably the most efficient medium for communication between humans. Optical character recognition has emerged as one of the most successful applications of technology in the artificial intelligence and the field of pattern recognition. Optical character recognition (OCR), is the process of converting handwritten text (numerals, letters) or scanned images of machine printed, into a computer format text. A Text-To-Speech (TTS) synthesizer is a system based on computer that should be able to read any text aloud, whether it was directly introduced in the computer by an operator or scanned and submitted to an Optical Character Recognition (OCR) system. The system consist of operational stages such as image capture, image preprocessing, image filtering, character recognition and text to speech conversion. The software platforms used are MATLAB, LabVIEW[1]. Along with the privilege of OCR technology the paper indulges modern sixth sense technology. Sixth Sense can also be a great substitute for many hardware devices. The sixth sense will recognize the finger movements .A person can resize and edit picture by making few finger motions. The submersion of these technologies will make operations easier for users.

## II . LITERATURE SURVEY

Ray Kurzweil started a company Kurzweil Computer Products, Inc. in 1974 and developed omni-font OCR. This omni-font OCR could recognise text printed virtually in any font. In the late 1960s and 1970s .Omni-font OCR was then used by companies, including CompuScan. Kurzweil then decided that the best application of this technology would be to create a reading machine for the blind in the late 1960s and 1970s. This technology made it possible for the computer to read text

aloud for blind people. This device further made the invention of two enabling technologies – the CCD flatbed scanner and the text-to-speech synthesiser. The of Sixth Sense technology prototype was developed by Steve Mann. This device is named Telepointer. It consists of hands-free, headwear-free device that facilitates the wearer to experience a visual collaborative telepresence, along with text, graphics, and a shared cursor which is displayed directly on real world objects Mann referred this device as “Synthetic Synesthesia of Sixth Sense”.

## III. PROPOSED SYSTEM

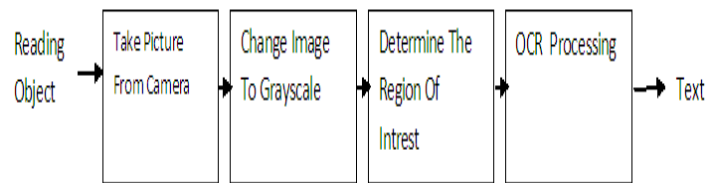


Fig.1. OCR System

**A. OCR SYSTEM:** Initially the system will create an OCR session which will be consisting of image detection i.e character detection through row wise and column wise scanning mechanisms. After successful detection of an image it will be first converted into its Grey form from RGB using the RGB thresholding concept [3]. Immediately after the conversion this system will compare the detected text or the particular character with the pre-loaded characters in their image form in a database which prepared according to system requirements. This will ensure a correct prediction of the detected character using the genetic algorithm[4]. After successful prediction of the whole word or the text given in the image it will be directly saved in the notepad. This notepad file will be then opened using the text to speech converter software which will convert the given file in speech

### B. Image Processing

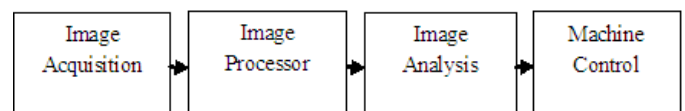


Fig.2. Image Processing Block

**Image acquisition setup:** It consists of suitable interface for connecting a web camera to PC.

**Processor:** It consists of personal computer or a dedicated image processing unit.

*Machine control:* After processing, some conclusions have to be made in order to initiate control actions. In this paper, control actions are made desktop control via mouse control.

*Image analysis:* Certain software tools are used to analyze the content in the image captured and derive conclusions e.g. Matlab 7[5].

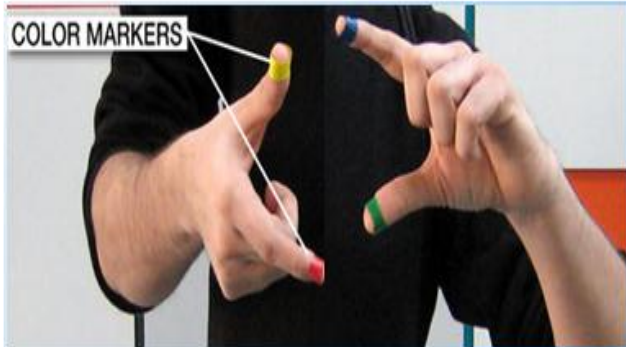


Fig.3. Color Markers

IV.METHODOLOGY

First stage is the image acquisition system, in which system consists of interfaced webcam, to capture the image of the text document. This image then goes through pre-processing, in which the region of interest (ROI) is obtained where-in, result obtained is separate sentences, and then words are separated and segmented. This data is then given to Template Identification, where the characters are detected, and individual alphabets are obtained. This data is then given to the OCR Algorithm which converts the image data to text data. For OCR a program is written for better outputs. The Algorithm scans the image, checks each alphabet or letter and gives a corresponding text output after verifying it with its own database[6].A Dictionary is used to compare the words detected by the Algorithm for auto-correction[7]. But this is optional. Next step is Storage Devices which can save the text data that is obtained after applying the algorithm in a text file. According to the application required the next step's function varies.Text to speech is chosen where the text data is converted to an audio output and is played through the earphones connected to the audio jack.

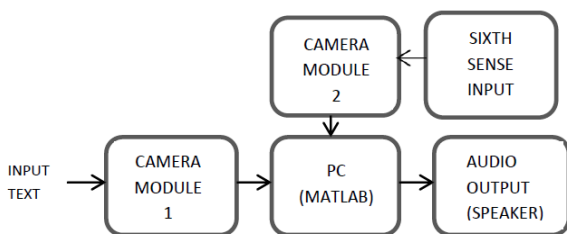


Fig.4. Construction of Main System

In the above shown block diagram of the proposed system an additional block is added which is the sixth sense input. This input is nothing but the input through RGB colour stickers stucked on to the user's fingers to provide some additional features like cropping the given image or for taking the snapshot of the particular area of interest [8]. This block is added to enhance the basic operations of the proposed system

and making it useful for various applications rather than for a specific purpose of text to speech conversion itself.

The *genetic algorithms* were first suggested , by John Holland in 1975 Currently using in a range of problems together with scheduling, images creating, planning strategy, predicting with dynamical systems, classification. For the application of genetic algorithm this study has used MATLAB 7.0.1 package to initialize application values.

The system has applied a set of binary numbers (1,0bits) representing English alphabets. The alphabets are from A to Z represented in a matrix which behaves as initiated value for comparing.

V.RESULT

Using camera module and Matlab

Input:



Input image

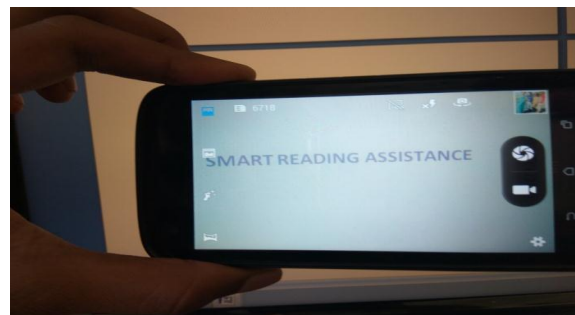
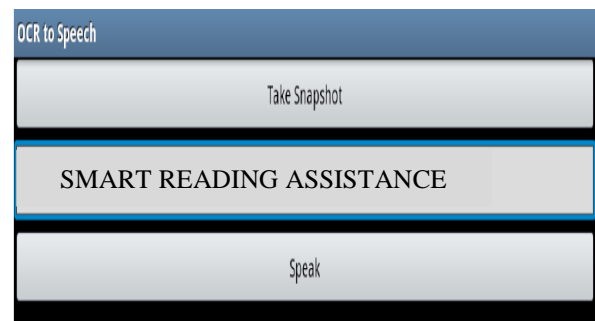


Image Capture

Output:



Output

VI.CONCLUSION

This paper is a small step towards helping a physically challenged people and lot more can be done to make the product more sophisticated, user friendly and efficient. People with poor vision or visual dyslexia or totally blindness can use this approach for reading the documents and books. People with speech loss or lazier people can utilize this approach to turn typed words into vocalization.

## REFERENCES

- [1] Christopher G Relf Image Acquisition and Processing with LabVIEW, CRC Press, 2004.
- [2] Smith, R. In proceedings of Document analysis and Recognition. An Overview of the Tesseract OCR Engine ICDAR 2007. IEEE Ninth International Conference.
- [3] J. Zhang and R. Kasturi, In IAPR Workshop on Document Analysis Systems "Extraction of Text Objects in Video Documents: Recent Progress," 2008.
- [4] N. Otsu, "A threshold selection method from gray-level histograms," In IEEE Trans on system, man and cybernetics, pp. 62-66, 1979.
- [5] Nikos Nikolaou , N. Papamarkos, Color reduction for complex document images, Int. Journal of Imaging Systems and Technology, v.19 n.1, p.14-26, March 2009
- [6] Huang, B.; Zhang, Y. and Kechadi, M.; Preprocessing Techniques for Online Handwriting Recognition, Springer Berlin Heidelberg, 2009, [doi>10.1002/ima.v19:1] Vol. 164
- [7] Tony F. Chan and Jackie (Jianhong) Shen (2005). Image Processing and Analysis - Variational and Stochastic Methods. Society of Industrial and Applied Mathematics. ISBN 0-89871-589-X.
- [8] C. Gonzalez and R. E. Woods, digital image processing, Pearson education, 2002.