VISUAL ASSIST

S Ashok[1], D Mahesh Kannan[2], N Jai Prasanth[3], Dr B Mathivanan[4]
[1, 2, 3, 4] Department of Computer Science and Engineering,
[1, 2, 3, 4] Sri Ramakrishna Engineering College, Coimbatore, INDIA

Abstract- *It is found that the visually challenged people use Braille system to read and learn which is difficult and complex process for them. Visual assist helps the visually impaired people in learning with ease. The main motive is to improve the learning capability of the visually challenged people in the upcoming years. Visual Assist as an application will use OCR to convert the image containing the text into a text document and TTS will be used to convert the text document into speech, so that the visually challenged people can hear and learn. By the advancement of the technology visually challenged people can easily learn when compared to the existing Braille system, where they had to read the text through the raised dots.*

Keywords- OCR-Optical Character Recognition, TTS-Text-to-Speech synthesizer.

I. INTRODUCTION

Image processing is a method to convert an image into digital form and perform some operations on it, in order to get an enhanced image or to extract some useful information from it. It is a type of signal dispensation in which input is image, like video frame or photograph and output may be image or characteristics associated with that image. Usually Image Processing system includes treating images as two dimensional signals while applying already set signal processing methods to them.

1. PROCESS OF IMAGE PROCESSING

In imaging science, image processing is processing of images using mathematical operations by using any form of signal processing for which the input is an image, a series of images, or a video, such as a photograph or video frame; the output of image processing may be either an image or a set of characteristics or parameters related to the image. Most image-processing techniques involve treating the image as a two-dimensional signal and applying standard signal-processing techniques to it. Images are also processed as three-dimensional signals where the third-dimension being time or the z-axis.

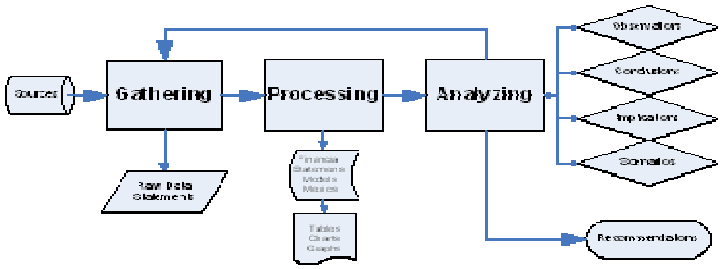

Figure 1. Overview of Image processing

Image Acquisition :

This is the first step or process of the fundamental steps of digital image processing. Image acquisition could be as simple as being given an image that is already in digital form. Generally, the image acquisition stage involves preprocessing, such as scaling etc.

Image Enhancement :

Image enhancement is among the simplest and most appealing areas of digital image processing. Basically, the idea behind enhancement techniques is to bring out detail that is obscured, or simply to highlight certain features of interest in an image. Such as, changing brightness & contrast etc.

Image Restoration :

Image restoration is an area that also deals with improving the appearance of an image. However, unlike enhancement, which is subjective, image restoration is objective, in the sense that restoration techniques tend to be based on mathematical or probabilistic models of image degradation.

Color Image Processing :

Color image processing is an area that has been gaining its importance because of the significant increase in the use of digital images over the Internet. This may include color modeling and processing in a digital domain etc.

Wavelets and Multiresolution Processing :

Wavelets are the foundation for representing images in various degrees of resolution. Images subdivision

successively into smaller regions for data compression and for pyramidal representation.

Compression :

Compression deals with techniques for reducing the storage required to save an image or the bandwidth to transmit it. Particularly in the uses of internet it is very much necessary to compress data

2. IMAGE TO SPEECH CONVERSION

Speech synthesis is the artificial production of human speech. A computer system used for this purpose is called a speech computer or speech synthesizer, and can be implemented in software or hardware products. A text-to-speech (TTS) system converts normal language text into speech; other systems render symbolic linguistic representations like phonetic transcriptions into speech.

Synthesized speech can be created by concatenating pieces of recorded speech that are stored in a database. Systems differ in the size of the stored speech units; a system that stores phones or diphones provides the largest output range, but may lack clarity. For specific usage domains, the storage of entire words or sentences allows for high-quality output. Alternatively, a synthesizer can incorporate a model of the vocal tract and other human voice characteristics to create a completely "synthetic" voice output.

The quality of a speech synthesizer is judged by its similarity to the human voice and by its ability to be understood clearly. An intelligible text-to-speech program allows people with visual impairments or reading disabilities to listen to written works on a home computer. Many computer operating systems have included speech synthesizers since the early 1990s.

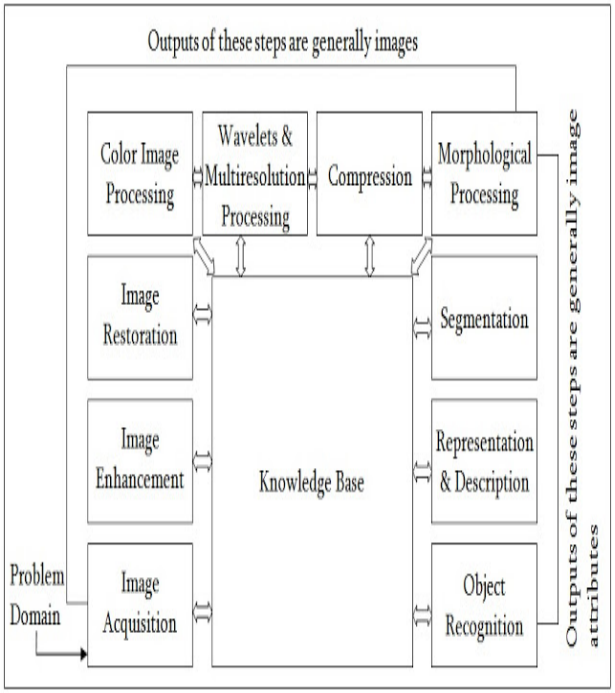

Figure 2. Architecture of Image processing

II. RELATED WORKS

The literature survey for this project focuses on the OCR processing and the related works for the future enhancements.

K Nirmala Kumari1and Meghana Reddy[1] proposed that the OCR conversion and Text to Speech can be carried out using Raspberry pi. But this paper fails to explain the portability. Archana A. Shinde[2] explains the mechanism of text preprocessing and the text segmentation for the OCR. R. Smith[3] . Emphasis is placed on aspects that are novel or at least unusual in an OCR engine, including in particular the line finding, features/classification methods, and the adaptive classifier.

III. EXISTING SYSTEM

In existing method, the Optical Character Recognition and text to speech synthesizer is implemented in the Raspberry Pi system which provided a greater use but it lacked the portability feature as it must have an external web camera. Thus the efficiency is reduced since the portability is external.

1. Disadvantage

- Requires external camera
- Less portable

2. Proposed method

In this paper, we describe by allowing the user to capture the text document as an image, using OCR image to text conversion is done and further the converted text is synthesized as voice using text to speech synthesizer. It helps the visually challenged people in reading and learning.

3. Advantage

- More portable
- Efficient performance

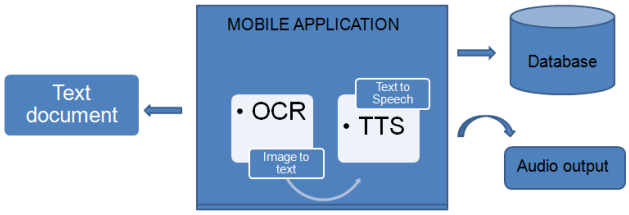

Figure 3. Architecture Diagram

IV. MODULE DESCRIPTION

In the first layer, the hard text document which is to be converted is being captured through the camera of the mobile device. Then in the second layer the captured image is processed and converted to fetch the text from the image using the Optical Character Recognition technique. Once the processing is completed then the processed text from the image is being stored in the database. MySQL is the database which is been used to store the processed text. Once the text is been stored then in the next layer the text is been converted to audio through the Text to Speech Synthesizer engine and a final audio is been given as the output through the speaker of the mobile device. This application also helps the user to make call to a phone number if there is a phone number mentioned in the image.

1. Image to Text Conversion

In this segment, the image is captured in the camera and it is scanned. The scanned image is processed by the OCR to identify the text characters in the scanned picture. The scanned text is stored in the SQL lite database. And the scanned text is forwarded to the text to speech module.

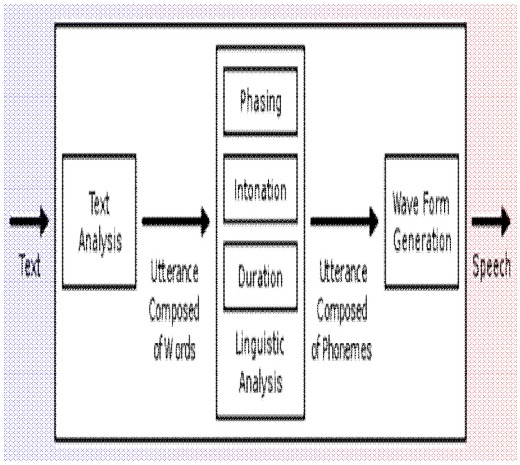

Figure 4. Module description

2. Text to Speech module

The scanned text is processed into the text to speech engine where the text format is converted into the audio file. Thus the converted audio can be accessed by the user through speakers or through ear phones.

V. CONCLUSION

The system was implemented on mobile device which is compact and easy to be carried to any place. The visual assist is a effective application for the visually impaired people which helps them to read any printed document in an audio format. The performance of the system is maintained in a proper way. The SQLite database which is used helps in a major way where it helps to store the information which is scanned.

REFERENCES

[1] K Nirmala Kumari1, Meghana Reddy J2 Associate Professor, Dept. of ECE, "Image Text to Speech Conversion Using OCR Technique in Raspberry Pi", BIT College, Bangalore, Karnataka, India PG Student [VLSI and E.S], Dept. of ECE, BIT College, Bangalore, Karnataka.

[2] Archana A Sindhe, "Text Pre-processing and Text Segmentation for OCR" D. International Journal of Computer Science Engineering and Technology 2012.

[3] R. Smith, "An Overview of the Tesseract OCR Engine" USA: Google Inc.