A JOURNEY THROUGH EXISTING PROCEDURES ON OCR RECOGNITION AND TEXT TO SPEECH CONVERSION PROCESS OF BILINGUAL LANGUAGES

K.Shanmugam[1], Dr.B.vanathi[2]
[1]Assistant Professor,[2]Professor,
[1,2]Department of Computer Science and Engineering
Valliammai Engineering College, Chennai,Tamilnadu

ABSTRACT

In this paper we try to explore various Character recognition and Text-to-Speech synthesis techniques developed and implemented by several researchers and research groups over the world during the last decades. Various existing proposals focus on the implementation of either OCR or Text to Speech synthesis. Text segmentation and recognition of Tamil characters poses a new challenge owing to its large number of characters, curves, indefinite shapes; techniques that recognise such text are discussed. This paper aims to give an overview of OCR and text synthesis for Bilingual languages (English and Tamil languages), summarizes and compares the characteristics of various synthesis techniques used.

KEYWORDS:

Optical character recognition, Text-To-Speech (TTS) synthesizer, Win32 SAPI.

INTRODUCTION:

Machine replication of human functions like reading, is an ancient dream. However, over the last five decades, machine reading has grown from dream to reality.

The two techniques that this paper discuss is about Optical Character Recognition (OCR) and Text To Speech (TTS) Synthesizer.

Optical character recognition has become one of the most successful applications of technology in the field of pattern recognition and artificial intelligence.Optical character recognition (OCR), is the process of converting scanned images of machine printed or handwritten text, into a computer format text.

The webcam is manually focused towards the text. Then, picture is taken and processed by Microcontroller to hear the spoken words of the text through the earphone or speaker plugged into it through its audio jack.

OCR is one of the challenging areas of character recognition and it has various practical applications.

A Text-To-Speech (TTS) synthesizer is a computer-based system 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 recognised text is converted into speech by using text to speech synthesizer.
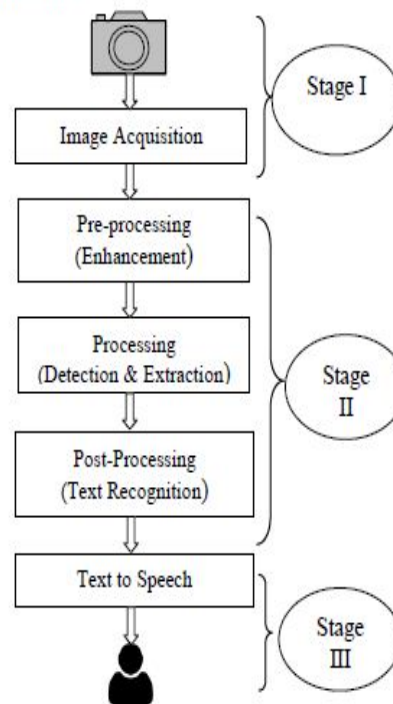

Figure 1: PROCESS OF IMAGE TEXT TO SPEECH [6]

OCR PROCESS:

Basic overview of how an OCR engine processes an image to text contained it: An image of the document is acquired by the computer.

1) An image of the document is acquired by the computer.

2) The image is submitted as input to an OCR engine.

3) The OCR engine matches portions of the image to shapes it is instructed to recognize.

4) Given logic parameters that the OCR engine has been instructed to use, the OCR engine will make its best guess as to which letter a shape represents.

5) OCR results are written as text.[1]

The OCR mainly consists of two processes. 1. Feature extraction 2.classification. The initial process involves the recognition of features of the character in the image then the classification is carried out.
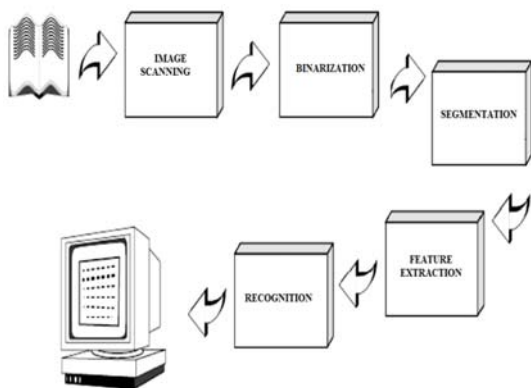

FIGURE 2: Components of an OCR system[2]

Features involve height, width, curves, size etc of the character. Classification is done by comparing the extracted features of the character with the pre defined libraries.

Libraries are created by training the datasets in neural network. After the feature extraction, the properties are given to a feed forward network, which in turn gives the character that matches the features.

TEXT TO SPEECH SYNTHESIZER:

A Text-To-Speech (TTS) synthesizer is a computer-based system that should be able to read any text aloud. A text-to-speech system 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 abbreviations 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 transcriptions to words is called text-to-phoneme conversion.

The back-end often referred to as the synthesizer—then converts the symbolic linguistic representation into sound.[2]
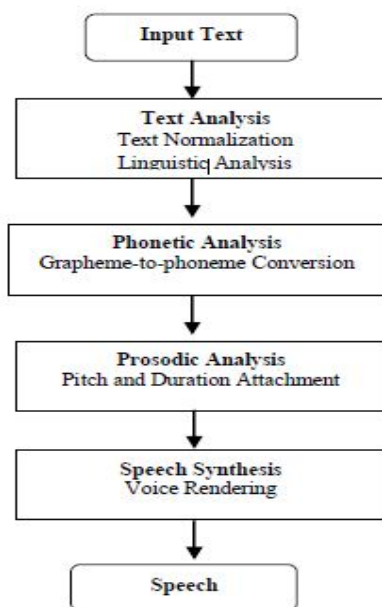

FIGURE 3: Block diagram of Text to speech Synthesis

RASPBERRY PI:

Raspberry Pi board is to run image recognition programs on it and a Headphone to deliver the output speech. The **Raspberry Pi** is a credit card sized single-board computer developed in the UK by the Raspberry Pi Foundation with the intention of stimulating the teaching of basic computer science in schools.[8]
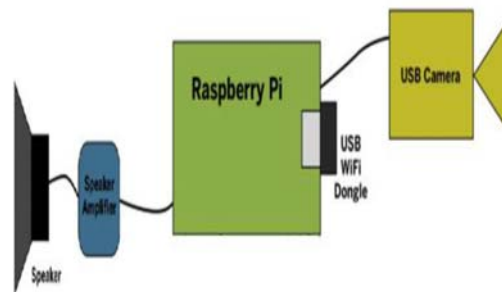

FIGURE 4: Block diagram[8]

TAMIL LANGUAGE:

Tamil is an Indian language spoken widely in TamilNadu in India.The character recognition process becomes difficult for some traditional Indian languages like Tamil. Since Tamil language consists of large characters it is a difficult task to feed all the characters into the libraries along with its properties. So classification is a big challenge.

The OCR process for Tamil language in various papers is analysed below:

Preprocessing: Preprocessing step involves noise removal using proper filters like mean filter, min-max filter, Gaussian filter. Binarization process converts gray scale or coloured image to black and white image[4].

Segmentation: After pre-processing, the noise free image is passed to the segmentation phase. At first level, lines are segmented using row histogram and words are extracted using column histogram and finally characters are extracted from words[4].

Feature extraction: Techniques like principle component analysis(PCA), Linear Discriminant Analysis(LDA), Independent Component analysis(ICA), Chain Code(CC), Scale Invariant Feature Extraction(SIFT) might be applied to extract the features of individual characters.[4]
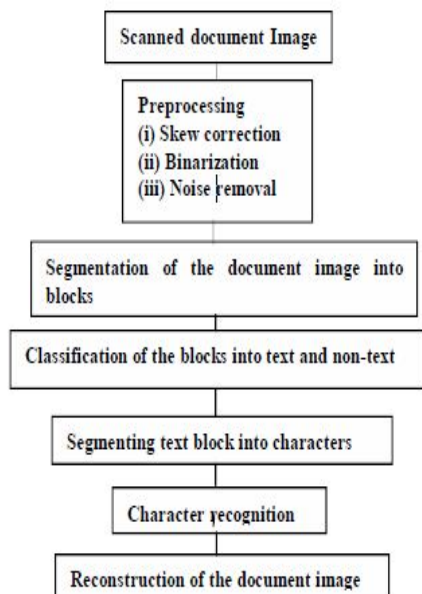
FIGURE 5: STEPS INVOLVED IN COMPLETE OCR PROCESS FOR TAMIL DOCUMENTS [3]

CLASSIFICATION: Classifiers like artificial neural networks or support vector machine[4] or a Radial basis function (RBF) neural network[3] are used.

TEXT TO SPEECH: Win32 SAPI library should be loaded in the computer. Gets the voice object from Win32 SAPI. Compares the input string with Win32 SAPI string.

Extracts voice by firstly select the voice which are available in library. Finally get the speech for given image[7].

REVIEW ON PAPERS:

TITLE	AUTHOR	PUBLISHED	TECHNIQUE USED	INFERENCE
Character Detection and Recognition System for Visually Impaired People	Akhilesh A. Panchal, Shrugal Varde, M.S. Panse	IEEE International Conference On Recent Trends In Electronics Information Communication Technology, May 20-21, 2016, India	Feature extraction algorithm applied on detected text using connected components (CC) analysis using MATLAB software	False detection of the scene image reduces accuracy of the system. Increases processing time and decreases speed of the algorithm
Text to Speech Conversion Using FLITE Algorithm	Tejashree M. Shinde, V. U. Deshmukh, P. K. Kadbe	International Journal of Science and Research (IJSR) ISSN (Online): 2319-7064,2013	Flite Algorithm	Can't make the time less than 1 second for the 20 word utterance
A Complete Tamil OCR system	Aparna.K.G,A.G. Ramakrishnan	White Paper	Bilinear Interpolation	Loss of Information in the thinned part of the image.
An OCR system for Tamil Newsprint	K.H.Aparna, Sumanth Jaganathan, P.Krishnan, V.S. Chakaravarthy	White Paper	RBF neural network	Document segmentation and character recognition is not accurate

CONCLUSION:

Techniques like optical character recognition (OCR) and speech synthesis proposed in various papers are analysed and brought together in this paper. In India some research organizations are also working on Text-to-Speech in regional languages like Marathi, Hindi, Telugu, Punjabi, Kannada etc. But all these systems do not

replicate the human natural speech. There is huge scope in improvement in the text to speech field to achieve a high level of natural and emotion aspect

REFERENCE:

[1]Optical Character Recognition for English and Tamil Script by *Kanimozhi.V.M, Muthumani. I* ,Alagappa chettiyar college of engineering & Technology, Karaikudi, India.

[2]Text to speech conversion system using OCR by *Jisha Gopinath, Aravind S, Pooja Chandran, Saranya S S*, SBCEW, Kerala, India

[3] An OCR system for Tamil Newsprint by *K.H.Aparna, SumanthJaganathan , P.Krishnan,V.S.Chakaravarthy*, Department of Electrical engineering,IIT Madras,Chennai.

[4]Tamil Handwritten Character Recognition :Progress and challenges by *K.Punitharaja,P.Elango*

[5] A Complete Tamil OCR system by *Aparna.K.G, A.G.Ramakrishnan,* Indian Institute of Science,Bangalore.

[6] Character Detection and Recognition System for Visually Impaired People by *Akhilesh A. Panchal, Shrugal Varde, M.S. Panse* .

[7]Implementation of Text to Speech Conversion by Chaw *Su Thu Thu, Theingi Zin,* Mandalay University,Mandalay.

[8]Embedded Optical Character Recognition On Tamil Text Image Using Raspberry Pi *V. Ajantha Devi, Dr. S Santhosh Baboo,*Department of Computer Science D.G.Vaishnav College,Chennai Tamilnadu, India