# NEW APPROACH FOR DEVNAGARI TEXT TO SPEECH CONVERSION

**Ms. Poonam S. Shetake[1], Dr. Mr. S.A.Patil[2], Mr. Prashant M. Jadhav[3]**
*[1]Department of E & TC, DKTE's YCP Polytechnic, Ichalkaranji*
*[2] Department of E & TC, TEI, Rajwada Ichalkaranji*
*[3] Department of Electronics, TEI, Rajwada Ichalkaranji*

## ABSTRACT

*A text to speech converter convert's normal language text into speech. Text to speech converter is useful in different applications. Customer support dialog systems Interactive voice response (IVR) systems etc and are also useful in an applied research. This application is more helpful in banking, toys and many other applications like checking marks, railways, aid to the physically challenged persons, language education and fundamental and applied research. etc. But text to speech conversion is not that much easy for machine as it is for human. Basic steps that machine has to follow for text to speech analysis are database creation, character recognition and text to speech conversion. For this paper we have used self created data base using susha font and for character recognition template matching algorithm is implemented. And according to the result of template matching related wave files are played.*

*Keywords: -segmentation, recognition, concatenating*

## INTRODUCTION

Language is the ability to express one's thoughts by means of a set of signs, whether graphical gestural, acoustic, or even musical. It is distinctive nature of human beings, who are the only creatures to use such a structured system. Speech is one of its main components. It is by far the oldest means of communication between human being and also the most widely used. No wonder, then, that people have extensively studied it and often tried to build machines to handle it in acoustic way. Most of the Information in digital world is accessible to a few who can read or understand a particular language.

Language technologies can provide solutions in the form of natural interfaces so the digital content can reach to the masses and facilitate the exchange of information across different people speaking different languages. These technologies play a crucial role in multi-lingual societies such as India which has about 1652 dialects/native languages. A text to speech converter convert's normal language text 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 diaphones provides the largest output

64

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.

Here question arises that whether machine or simply computer can perform same task of text to speech conversion? Answer is not that much easily as human can. The machine has to follow some procedure which is divided in basic two steps. At first character recognition for this template matching method is used. Next step is TTS that is Text to speech conversion in this we have to convert recognized text from template matching into .wav file or simply in speech file.

Text to speech conversion is still challenging problem because of some external factors like while recording .wav files noise may get added in it and also while preparing database for characters it should consist all images of same dimensions & size.
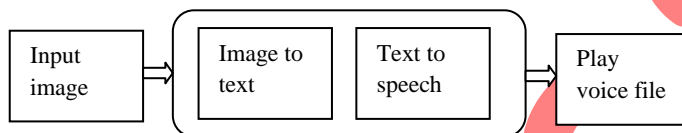
Fig.1 Block diagram of system

## PROPOSED TTS CONVERSION METHOD

This section describes TTS conversion method. The proposed method consists of four parts: database reparation, pre-processing, character recognition using template matching algorithm, speech conversion by playing recognized character using the pre-stored voice templates. Each of them is described as follows:
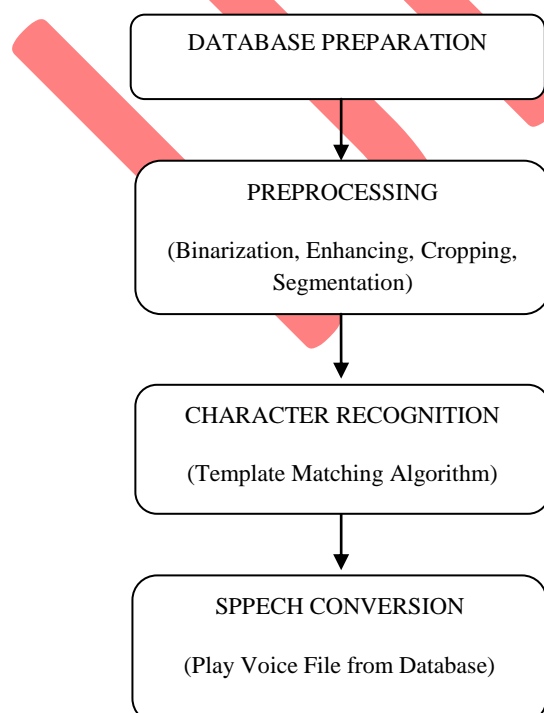
Fig. 2 TTS conversion system

## A. DATABASE CREATION

For this implemented paper two data bases are used first for characters or templates and second is wave file database.

1. Character database :

For creation of this database shusha font is used for that character is pasted in a paint where it is cropped and saved as a jpg .image file.

2. Wave file database:

For this creation of this database, recording of all consonants and vowels s carried out in Sangli studio continuous recording of all alphabets is done by using their studio so as to maintain minimum addition of noise. Then by using sound fold software total recorded file is sampled into different parts and saved according to respected alphabet.

## B. IMAGE PREPROCESSING

Images saved in character database are may be affected because of background, light effects if downloaded from internet and also due to variation in size . Hence preprocessing is required. The steps followed in preprocessing are:

1. Grayscale conversion / binarization : As database consist color photographs there is need to convert RGB color values to gray color using MATLAB. Then these gray colored images are enhanced in next step. Grayscale images is result of measuring intensity of each pixel. For

achieving accuracy input document should be grayscaled. To convert a colour from a colourspace based on an RGB colour model to a grayscale representation following function is used

$Y = 0.2126R + 0.7152G + 0.0722B$

नमन

Fig 3. Image before grascale conversion

नमन

Fig 4. Image after grascale conversion

2. Cropping: here image is cropped using top , left and right, bottom to remove line. Also it is used to remove some extra portion from top.

नमन

Fig.5 Image before cropping

नमन

Fig 6 Image after cropping

3. Thresholding :

नमन

Fig 7 Image after Thresholding

4. Vertical & horizontal Segmentation:

न    म    न

Fig 8 Segmented characters

## C.  CHARACTER RECOGNITION

Optical Character Recognition by using Template Matching is a system that is useful to recognize the character or alphabet by comparing two images of the alphabet. Template Matching is the algorithm is used here to recognize the characters. Characters to be tested are in the devanagari form.

Basically it consist jpg image of some standard size of 240 x 240 which is converted into grayscale in previous step. This algorithm compares two alphabets. Template matching involves determining similarities between a given template and windows of the same size in an image and identifying the window that produces the highest similarity measure. It works by comparing derived image features of
the image and the template for each possible displacement of the template.

Steps used in template matching algorithm are as follows:
i) Firstly, the character image from the detected string is selected.
ii)  After that, the image to the size of the first template is rescaled.
iii)  After rescale the image to the size of the first template (original) image, the matching metric is computed.
 iv)  Then the highest match found is stored. If the image is not match repeat again the third step.
v)  The index of the best match is stored as the recognized character.

67

### D. SPEECH CONVERSION

Speech is the vocalization form of human communication. Speech communication is more effective medium than text communication medium in many real world applications.

Speech synthesis is the artificial production of human speech. A system used for this purpose is called a speech synthesizer, and can be implemented in software or hardware. Synthesized speech can be created by concatenating pieces of recorded speech that are stored in a database.

## SYSTEM IMPLEMENTATION AND RESULTS

A text to speech [TTS] system is a system that can convert a given text into speech signals. The aim of an ideal TTS system is to be able to process any text that a human can read. A TTS system consists of mainly two parts: Text processing part and speech synthesis part. Here the output of a recognized character system is used as an input for this system.



```
ufname =

k.wav


ufname =

p.wav
```

Fig 9 File names searched to play word

## CONCLUSION & FUTURE SCOPE:

This paper is an effort to suggest an approach for image to speech conversion using character recognition and text to speech technology. The application is user friendly, cost effective and applicable in the real time. By this approach we can read text from a document. System is very useful for blind persons to read the documents.

Experiments have been performed to test the text reading system and good results have been achieved.

68

## REFERENCES

[1]Srikrishna,  "*Text  Extraction  From  Etrogenous  Imagesusing  Mathematical Morphology*",Journal of Theoretical and Applied Information Technology 2005-2010.

[2]   M. Nagamani, S.Manoj Kumar, S.Uday Bhaskar, " *Image to Speech Conversion System for Telugu Language*", International Journal of Engineering Science and Innovative Technology (IJESIT) Volume 2, Issue 6, November 2013.

[3] Alexandre Trilla, "*Natural Language  Processing techniques in Text-To-Speech  synthesis and Automatic Speech Recognition*",Departament de Tecnologies M`edia Enginyeria i Arquitectura La Salle (Universitat Ramon Llull), Barcelona, Spain2009.

[4]   D.Sasirekha, E.Chandra, "*TEXT TO SPEECH: A SIMPLE TUTORIAL*"International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Volume-2, Issue-1,March 2012.

[5]   Bhushan Sonawane,Kiran Patil, Nikhil Pathak, Ram Gamane, "Third Eye : An Image Explorer",  International Journal of Emerging Technology  and Advanced Engineering Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 3, Issue 4, April 2013)

[6]   Priyanka Jose , Govindaru V, "  Malayalam Text-to-Speech"International Journal of Engineering and Technical Research (IJETR) ISSN: 2321-0869, Volume-1, Issue-3, May 2013.

[7]   Adam Coates, Blake Carpenter, Carl Case, Sanjeev Satheesh, Bipin Suresh, Tao Wang, David J. Wu, Andrew Y. Ng, " Text Detection and Character Recognition in Scene Images with Unsupervised Feature Learning",Computer Science Department Stanford University USA

[8]   A. Chauhan, Vineet Chauhan, Surendra P. Singh, Ajay K. Tomar, 5Himanshu Chauhan, "A Text to Speech System for Hindi using English Language", 322 International Journal of Computer Science and Technology IJCST Vol. 2, Issue 3, September 2011.

[9]   Adil Farooq, Ahmad Khalil Khan, Gulistan Raja, "Implementation of a Speech Based Interface System for Visually Impaired Persons",Department of Electrical Engineering, UET Taxila, Pakistan. Life  Science Journal 2013.

[10]   Orhan Karaali, Gerald Corrigan, Noel Massey, Corey Miller, Otto Schnurr and Andrew Mackie, "  A HIGH QUALITY TEXT-TO-SPEECH SYSTEM  COMPOSED OF MULTIPLE NEURAL NETWORKS",  IEEE International Conference on Acoustics, Speech and Signal Processing, Invited paper, Seattle, May 1998.

[11]   Julinda Gllavata1, Ralph Ewerth1 and Bernd Freisleben, "A Robust Algorithm for Text Detection in Images"

[12]   Prof. Amit Choksi1, Nihar Desai2, Ajay Chauhan3, Vishal Revdiwala4, Prof. Kaushal Patel5, "Text Extraction from Natural Scene Images using Prewitt Edge Detection Method", International Journal of Advanced Research inComputer Science and Software Engineering Volume 3, Issue 12, December 2013 ISSN: 2277 128X.

[13]   Tapas Kumar Patra , Biplab  Patra ,Puspanjali Mohapatra "Text to Speech Conversion with Phonematic Concatenation"International Journal of Electronics Communication and Computer Technology (IJECCT) Volume 2 Issue 5 (September 2012) ISSN:2249-7838 .

[14]   Pijus KASPARAITIS , "Diphone  Databases for Lithuanian Text-to-Speech  Synthesis", INFORMATICA, 2005, Vol. 16, No. 2, 193–202 193,  2005  Institute of Mathematics and Informatics, Vilnius.

[15]   K. Partha Sarathy, A.G.Ramakrishnan, "TEXT TO SPEECH SYNTHESIS SYSTEM FOR MOBILE APPLICATIONS",Department of Electrical Engineering, Indian Institute of Science, Bangalore, India.

[16]   Prof. Sheetal A. Nirve,Dr. G. S. Sable,"Optical character recognition for printed text in Devanagari using ANFIS",International Journal of Scientific & Engineering Research, Volume 4, Issue 10, October-2013 236 ISSN 2229-5518 IJSER © 2013.