



INTERNATIONAL JOURNAL OF PURE AND APPLIED RESEARCH IN ENGINEERING AND TECHNOLOGY

A PATH FOR HORIZING YOUR INNOVATIVE WORK

REVIEW OF IMAGE TEXT TO SPEECH CONVERTER

D. V. MESHAM, VAISHNAVI KAWDE, SHASHANK LANDOLE, DIVYANI MANE, NIKITA MARTAWAR, DIPALI HARSHE

Department of CE, SGBAU University, Amravati.

Accepted Date: 15/03/2016; Published Date: 01/05/2016

Abstract: This paper contains the functionality of Optical Character Recognizer [OCR] and speech synthesizer. Image is input for OCR and OCR extract the text form that image and converts it into machine language. This system can be useful in various applications like legal industry, banking, other industries, office and home automation. It is designed for people who are not able to read any type of text documents. In this paper, the character recognition method is presented by using OCR technology.

Keywords: Artificial intelligence, Optical character recognition, Speech synthesis.



PAPER-QR CODE

Corresponding Author: MR. D. V. MESHAM

Access Online On:

www.ijpret.com

How to Cite This Article:

D. V. Meshram, IJPRET, 2016; Volume 4(9): 1464-1468

INTRODUCTION

Language is the way to express one's thoughts by means of a set of signs, whether graphical gestural, acoustic. It is nature of human beings, who are the only creatures to use such a structured system. Speech is one of its main part. It is by far the oldest means of communication between human being and also the most widely used. No doubt, then, that people have studied it and often tried to build machines to handle it in acoustic way.

In this paper, we are extending this speech capability by developing a software which would provide a voice output for which we are using OCR. OCR (Optical Character Recognition) also referred to as Optical Character Reader is a system that translate printed character into alphanumerical form with electronic speed. Recognition engine of the OCR system translate the images of handwritten or printed characters into American Standard Code for Information Interchange (ASCII) data which are machine-readable characters. Therefore, OCR allows users to automate data capture from forms and still maintain the high level of accuracy and eliminate keystrokes to reduce data entry costs required in forms processing applications.

A text to speech (TTS) synthesizer is a windows based system that can read text automatically. A speech synthesizer can be developed by both software and hardware .

III. EXISTING SYSTEMS

In existing system Optical Character Recognizer[OCR] recognizes and extracts the character from bitmap image in the form of ASCII format. Output of character recognizer is acts as an input to the speech synthesizer.

A Text-To-Speech (TTS) synthesizer is a windows based system that should be able to read text a loudly, when it is directly given to the computer as input by an user. It is suitable to define Text-To-Speech or speech synthesis as an automatic generation of speech, by 'grapheme to phoneme' transcription.

1.Orhan Karaali, , Noel Massey, Gerald Corrigan Corey Miller, Otto Schnurr and Andrew Mackie proposed system for speech synthesis which consist Vocoder: The phonetic neural network is not trained to generate speech directly. Instead, it is trained to produce a sequence of auditory descriptions of 10 millisecond frames of speech. These are then synthesized using a vocoder.

2. Tapas Kumar, Patra Biplab Patra, Puspanjali Mohapatra states, the approach used here is a concatenative one. Today's high quality speech synthesizers are concatenative synthesizers. In a system, a person records speech containing a huge set of basic sound units, usually corresponding to a short sequence of phonemes.

3. Prof. Sheetal A. Nirve, Dr. G. S. Sable used Recognition algorithms like Neural Network. Neural network is also called as Artificial Neural Network (ANN). It is an artificial intelligent system which is based on biological neural network. Neural networks are capable for training a particular function by adjusting the values of the connections (weight) between these elements.

III. PROPOSED SYSTEM

In our proposed system, image is input for OCR and OCR extract the text from that image and converts it into machine language i.e. ASCII. Unlike existing system, our system will take different images as input such as JPG, GIF, RTS, etc. and it will be able to extract the text from it. That extracted text will act as input for speech synthesizer. Speech synthesizer is used to translate extracted text into the voice.

- **Optical Character Recognizer:**

Components of an OCR system :

1. Image Scanning, 2. Binarization, 3. Segmentation, 4. Feature extraction, 5. Recognition.

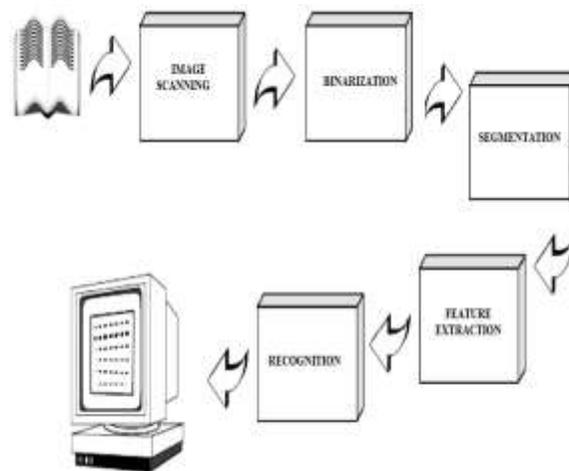


Fig: Components of OCR

Image scanning: Scanning makes original document as digital image. Scanning comes with thresholding. In Thresholding process multi level image is converted into bi-level image i.e. black and white image.

Binarization & Segmentation: Segmentation extracts text from figures and graphics. When segmentation is applied to text, it isolates characters or words.

Feature extraction & Recognition: It extracts the features of symbols. Features are the characteristics. The character image is converted to a higher level by extracting special characteristics of the image in the feature extraction phase. Recognition extracts text from images of documents.

- **Speech Synthesizer:**

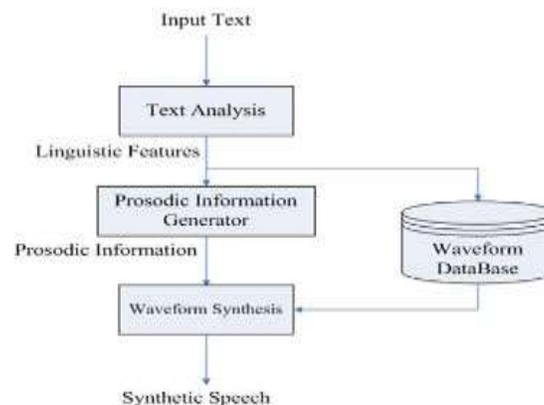


Fig. 1 The TTS system architecture

Text Analyser: The Text Analysis is part of pre-processing. The input text is analyzed and organized into list of words. Then it translates them into full text. Text Normalization is the conversion of text to pronounceable form. Text normalization is performed before text is processed in some way, such as generating synthesized speech.

Prosodic Information Generator:

The prosody generation system is consist of two main steps:

- An abstract symbolic description of the prosody is firstly generated from its syntax;

- This symbolic string is then translated into an acoustic description of prosody: phonemic duration and curve.

IV. CONCLUSION

The proposed centralized system for image text to speech converter allows the illiterate users to read the document . This paper made a clear and simple overview of working of OCR system and text to speech system (TTS) in step by step process. In the market , there are various text to speech systems (TTS) available and also much improvisation is under working in the research area to make the speech natural and more effective with stress and emotions.

V. REFERENCES

1. M.Ostendorf and I.Bulyko, "The impact of speech recognition on speech synthesis", in proc, IEEE Workshop Speech Synthesis, Santa Monica, 2002, pp. 99-106.
2. Text To Speech Synthesis - a knol by Jaibatrik Dutta.
3. K. Partha Sarathy, A. G. Ramakrishnan, "TEXT TO SPEECH SYNTHESIS SYSTEM FOR MOBILE APPLICATIONS", Department of Electrical Engineering, Indian Institute of Science, Bangalore, India.[CE] ISSN: 2231-2307, Volume-2, Issue-1, March 2012.
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. A.C. Bovic, M. Clark and W.S. Geisler, Multichannel texture analysis using localized spatial filters, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 12, n°1, pp 55-73, 1990
6. T.N. Tan and A.G. Constantinides, Texture Analysis based on a Human Visual Model, Proc. IEEE Int. Conf. On Acoust., Speech, Signal Proc., pp 2091-2110, Albuquerque, New Mexico, 1990
7. D. Dunn and al., Texture Segmentation Using 2-D Gabor Elementary Functions, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 16, n°2, 1994
8. Trier, Text, Evaluation of binarization methods for document images, IEEE Transactions on Pattern Analysis and Machine Intelligence, 1994.