

# A REVIEW ON TEXT RECOGNITION FOR VISUALLY BLIND PEOPLE

Siva Ganga s<sup>1</sup>, Hema s<sup>2</sup>

<sup>1</sup>M. Tech Student, Dept. of ECE, LBS Institute of Technology for Women, Kerala, India

<sup>2</sup>Assistant Professor, Dept. of ECE, LBS Institute of Technology for Women, Kerala, India

\*\*\*

**Abstract** - Nowadays text recognition is an essential problem for visually impaired peoples. Text-to-speech (TTS) is the generation of synthesized speech from the text. This paper presents a survey on text detection and the conversion of text to speech. There have been many methods for detecting text in past work. This paper reviews the techniques of different methods for text recognition.

**Key words:** Text-to-speech (TTS), Text recognition

## 1. INTRODUCTION

The number of visually impaired persons increased every year due to eye diseases, uncontrolled diabetes, accidents and other reasons. For a visually impaired person the most significant difficulty is to read. The people who are blind or have low visual impairment they have an ability to read printed labels and documents will enhance independent living. Reading is essential for every human being. In everywhere there is printed text which is in the form of reports, receipts, bank statements, restaurant menus, product packages, medicine bottles etc. can help blind users and those with low vision to read text, there are few devices that can provide good access to common handheld objects such as product packages, and objects printed with the text such as prescription medication bottles. To extract the text information from the camera captured image have different text patterns and different backgrounds. A novel feature is defining for extraction of text from the image. From natural scene images, it is very difficult to detect the text. The natural scenes contain different orientations shapes, sizes, and complex backgrounds. The ability of people who are blind or have significant visual impairments to read printed labels and product packages will enhance independent living.

## 2. RELATED WORK

As per the survey it has been observe there are enough amount of system has proposed to improve the difficult life of unsighted people. In [1], Rajput, Rajendrasing and Rushikesh Borse proposed a system which is used read text on product packaging from hand-held object for blind persons in their everyday life. Here Motion based method is used to detect the object of interest from background or other object in the camera view and then to extract text regions from complex backgrounds, text localization algorithm is used.

In [2], Sonal I Shirke, Swati V Patil proposes a portable camera-based assistive text reading framework which helps to the blind person to read text labels and product packaging from hand-held objects. The system framework consists of three functional components. At First, the scene capture-using a mini camera, and the text which the user needs to read gets captured as an image and sent it to the image or data processing platform. Then second, the data processing where text will be filtered from the surrounding and it will be recognized by optical character recognition (OCR) software, then finally, Speech output. A filtered text should be passed into this system to get an audio output. In addition the text recognition using OCR, this paper also includes template matching technique as a separate method for recognizing certain objects like currency notes.

In [3], Hao Jiang, Thomas Gonnot, Won-Jae Yi and Jafar Sanie propose a work which explained the decomposition process of text recognition and the Text-to-Speech algorithm. The text recognition process decomposes in three main steps. First the image is processed by the OCR algorithm, that inputs the region containing the sign with the text to be recognized, and returns a raw text. Then the text is forwarded to a second algorithm in charge of correcting the text and recovering sentences when possible. Finally, the result is converted to text using the Text-to-Speech algorithm. The user feedback is also composed of three main steps. Detecting the signs in the image is the first step. Once all the signs are located, they are forwarded to the OCR process. Then to optimize the recognition of the signs, another algorithm that will evaluate how the camera should be moved. Finally, generated a set of commands and sent to the Text to-Speech algorithm to be communicated to the user.

In [4], Adiga, Nagaraj, and S. R. Mahadeva Prasanna proposed the hybrid text-to-speech based on sub-band approach. To develop high-quality Text-to-Speech (TTS) this paper proposes a sub-band speech synthesis approach. For the low frequency band Hidden Markov Model (HMM)-based speech synthesis is used and high-frequency band waveform-based speech synthesis is used. Both of these methods show good performance and also it has benefits and short coming. Among one of the motivation is to apply the right speech synthesis method in the right frequency band.

In [5], Samruddhi Deshpande and Ms revati shriram propose a work which presents camera based system

which can help blind persons for reading text patterns printed on hand held objects. This framework is to assist visually impaired persons to read text patterns and then convert it into the audio output. The system proposed a method to capture the image from the camera and detect the object from background and then extract the text pattern from that object. Using Maximally Stable External Regions (MSER) feature detect text which are maximally stable. On the basis of variety of scenes a novel algorithm is evaluated. Compare the detected text with the template and then converted in to the speech output. Using the Optical Character Recognition (OCR) the text patterns are localized and binarized. The recognized text is converted to an audio output and the speech output is given to the blind user.

In [6], Granell, Emilio, Verónica Romero, and Carlos -D. Martínez -Hinarejos proposed the use of multimodal combination techniques for improving the CATTI system. Multimodal combination techniques have benefits of using speech as an additional source of information for the assisted transcription of historical manuscripts. The use of lattice combination techniques permitted to obtain the transcription outputs with a reduced error. This error reduction is only because the combination may produce new bigrams that increase the search alternatives, and that the adjustment of the word posterior probabilities should be increased the probabilities of the correct words. The main advantage of the presented approach is the error reduction produced by lattice combination techniques which allow reducing significantly the human effort when an assistive transcription system is used.

### 3. CONCLUSION

This review focuses on various techniques for implementing a text recognition system for visually blind or impaired people. This also briefs the different methods used for the system.

### REFERENCES

- [1] Rajput, Rajendrasing, and Rushikesh Borse. "Alternative Product Label Reading and Speech Conversion: An Aid for Blind Person." International Conference on Computing, Communication, Control and Automation (ICCUBEA). IEEE, 2017.
- [2] Sonal I Shirke 1, Swati V Patil. "Portable Camera Based Text Reading of Objects for Blind Persons." International Journal of Applied Engineering Research ISSN 0973-4562 Volume 13, Number 17 (2018) pp. 12995-12999
- [3] Jiang, Hao, et al. "Computer vision and text recognition for assisting visually impaired people using Android smartphone." 2017 IEEE International Conference on Electro Information Technology (EIT). IEEE, 2017.

- [4] Adiga, Nagaraj, and S. R. MahadevaPrasanna. "A hybrid Text-to-Speech synthesis using vowel and non vowel like regions." In India Conference (INDICON), 2014 Annual IEEE, pp. 1-5. IEEE, 2014.

- [5] Deshpande, Samruddhi, and Revati Shriram. "Real time text detection and recognition on hand held objects to assist blind people." International Conference on Automatic Control and Dynamic Optimization Techniques (ICACDOT). IEEE, 2016.

- [6] Granell, Emilio, Verónica Romero, and Carlos -D. Martínez -Hinarejos. "Image-speech combination for interactive computer assisted transcription of handwritten documents." Computer Vision and Image Understanding : 74-83 © 2019 Elsevier Inc.