SYNOPSIS

Report on

Voco Buddy

[Text To Speech]

by

Tanu Sharma 2000290140125
Rajat Deol 2000290140096
Himanshu Tomar 2000290140051

Session:2021-2022 (4th Semester)

Under the supervision of

Prof. Dr. Akash Rajak

KIET Group of Institutions, Delhi-NCR, Ghaziabad

DEPARTMENT OF COMPUTER APPLICATIONS
KIET GROUP OF INSTITUTIONS, DELHI-NCR, GHAZIABAD-201206
(MARCH- 2022)

ABSTRACT

There are about millions of blind and visually impaired people worldwide. These people are not able to see and hear anything. Disability of visual text reading has a huge impact on the quality of life for visually disabled people. Though many several devices have been designed for helping visually disabled to see objects using an alternating sense such as sound and touch, the development of text reading device is still at an early stage. Some existing systems for text recognition are typically limited either or require user assistance or may be of high cost. Therefore, some low-cost system must be developed that will enable to automatically locate and read the text aloud to visually impaired persons. In this paper we will study various ideas to recognize the text character and convert it into speech signal along with some applications of TTS systems. This paper also contains the various terms and concepts of text to speech conversion systems.

TABLE OF CONTENTS

INTRODUCTION

Text-to-speech (TTS) is a type of assistive technology that reads digital text aloud. It's sometimes called "read aloud" technology. With a click of a button or the touch of a finger, TTS can take words on a computer or other digital device and convert them into audio. TTS is very helpful for kids and adults who struggle with reading. But it can also help with writing and editing, and even with focusing. TTS works with nearly every personal digital device, including computers, smartphones, and tablets. All kinds of text files can be read aloud, including Word and Pages documents. Even online web pages can be read aloud. The voice in TTS is computer generated and reading speed can usually be sped up or slowed down. Many TTS tools highlight words as they are read aloud. This allows the user to see text and hear it at the same time. Some TTS tools can also read text aloud from images. For example, a user could take a photo of a street sign on their phone and have the words on the sign turned into audio. You might be wondering what the connection is between TTS and audiobooks. TTS is a tool that reads text aloud. An audiobook is a recording of a book read by a human voice (or created by TTS). Sometimes, people say TTS or audiobooks to mean the same thing. There are many different TTS tools: Built-in text-to-speech: Many devices have built-in TTS tools. This includes desktop and laptop computers, smartphones, digital tablets, and Chromebooks. Web-based tools: Some websites have TTS tools on-site. Text-to-speech apps: Users can download TTS apps on smartphones and digital tablets. There are also TTS tools that can be added to web browsers, like Chrome. Text-to-speech software programs: Many literacies software programs for desktop and laptop computers have TTS.

LITERATURE REVIEW

The Paper “Text extraction from documents images using edge information” focuses on the detection of text from a coloured images using edge-based feature. Its aim is to differentiate text from non-textual portion. However, there are some limitations like the generalized text segmentation method is not defined.

In this literature review, researcher has focused on the effectiveness of Text to speech technologies such as computer, smartphone, and tablets as assistive technology for blind students in the primary schools. The major objective of the scholar is to increase the academic performance of visual impairment children by use of Text to speech technologies. For this objective, scholar has discussed various driving forces such as technologies and its application. In this literature review, the effectiveness and practical application of the text speech technology has been addressed. With help of this literature review, the detail knowledge and information about the text to speech technology has been collected by the scholar because they used various articles of the different authors who have given the statement about the assistive technology. The major objective of this literature review is to conduct study on the topic of text to speech technology and its importance for the blind children who have visual impairment.

The paper “Text Recognition using Image Processing” focuses on text extraction using image processing. The image processing includes the pre-processing, segmentation, classification and post processing stages. However, there are some limitations like the model works offline, so if a user needs to perform text extraction, he/she need to perform the text extraction on their own system.

PROJECT OBJECTIVE

The main objective with this project is to help the persons who have learning disabilities or are visually challenged. The proposed system will prevent eye from strain ad user can sit and listen comfortably. The proposed system will save time especially while driving or exercising. The proposed system is easy to use and also helps in improving, reading and writing skills. Thus, making the entire process more secure, reliable, fast, and privacy oriented.

PROJECT METHODOLOGY

1. Referring project on extraction of Text from image into Speech.
2. The current system is referring to not natural sounding.
3. In the current system they are unable to read the symbols.

PROJECT OUTCOME

Proposed research work is elaborated using Python to calculate the efficiency of proposed techniques with existing techniques in terms of

- Accuracy
- Severity
- Specificity

PROPOSED TIME DURATION

The estimated time for the completion of this project is of 2 months.

		Week 0	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8
Feasibility Study										
Literature Review										
Implementation/Simulation										
Research Paper										
Report Writing										

REFERENCES

[1] Andersson S., Georgila K., Traum D., Aylett M., Clark R.A.J., "Prediction and Realisation of Conversational Characteristics by Utilising Spontaneous Speech for Unit Selection", 5th International Conference on Speech Prosody (Speech Prosody 2010), Chicago 2010, 1 – 2.

[2] Aylett M.P., Yamagishi J., "Combining Statistical Parameteric Speech Synthesis and Unit-Selection for Automatic Voice Cloning", Proceedings of LangTech, Rome 2008, 3.

[3] Aylett M. P., Potard B., Pidcock Ch.J., "Expressive speech synthe-sis: synthesising ambiguity", 8th ISCA Workshop on Speech Syn-thesis (SSW-8), ISCA, Barcelona 2013, 217.

[4] Balyan A., Agrawal S.S., Dev A., "Speech Synthesis: A Review", International Journal of Engineering Research and Technology IJERT, Vol. 2, Issue 6, 2013, 57 – 75.

[5] Bellegarda J.R., "Toward Naturally Expressive Speech Synthesis: Data-Driven Emotion Detection Using Latent Affective Analysis", 7th ISCA Workshop on Speech Synthesis (SSW-7), ISCA, Kyoto 2010, 200.

[6] Chandra E., Akila A., "An Overview of Speech Recognition and Speech Synthesis Algorithms", International Journal of Computer Technology and Applications, Vol.3, Issue 4, 2012, 1427.

[7] Chauhan A., Chauhan V., Singh G., Choudhary C., Arya P., "De-sign and Development of a Text-To-Speech Synthesizer System", International Journal of Electronics and Communication Technol-ogy, Vol. 2, Issue 3, 2011, 42 – 44.