NOVATEUR PUBLICATIONS
INTERNATIONAL JOURNAL OF INNOVATIONS IN ENGINEERING RESEARCH AND TECHNOLOGY [IJIERT]
ISSN: 2394-3696
VOLUME 7, ISSUE 6, June-2020

BOOK READER

VIRAJ KAMTEKAR
Student, Department of Computer Engineering
Vidyalankar Institute of Technology, Mumbai, India
viraj.kamtekar@vit.edu.in

OMKAR DESHMUKH
Student, Department of Computer Engineering
Vidyalankar Institute of Technology, Mumbai, India
omkar.deshmukh@vit.edu.in

VINIT RAUT
Student, Department of Computer Engineering
Vidyalankar Institute of Technology, Mumbai, India
vinit.raut@vit.edu.in

MOHINI CHAUDHARI
Assistant Professor, Department of Computer Engineering
Vidyalankar Institute of Technology, Mumbai, India
mohini.chaudhari@vit.edu.in

ABSTRACT
As stated by the statistics of the WHO-World Health Organization, 285 million people around the world are estimated to be visually impaired, of which 39 million are complete blind, out of which 90% live in developing countries. People who suffer from eyesight and visual defect are not capable of reading text in common newsprint, books, magazines and other reading materials clearly. This makes the reading process difficult which can disturb the learning process and slow the person's intellectual development. So, there is a need to develop a book reader device that is affordable to the low-income sections of the society. Thus, the objective of the system is to read data from the given input document and convert it into audio format using TTS technology. Hence, the system would take care that the student's disability won't be the major reason to hold them from pursuing their career.

INDEX TERMS: TTS-Text to Speech, OCR- Optical Character Recognition, Open CV- Open source computer vision, NoSQL- Non-Structured Query Language WHO- World Health Organization, Graphemes, and Phoneme

INTRODUCTION
Computers can perform the operations based on given instructions and the users can read the result of the same on the monitors or by printing the output. But there are situations where people need vocal communication with the computers. The TTS is a system for conversion of text documents to audio. One of the main applications of a TTS system is to help people with visual impairment and making them understand the textual content present on the computer.
A survey performed at local blind school in Mumbai stated that special aids and methods are required to make visual impaired students understand the concepts of subjects which are the part of their curriculum. Constant help and support are needed which makes the learning process much passive and very difficult for such students. Revising subjects for examination is also a major problem faced by the students. A full proof and user-friendly solution for this problem is to develop a web application which converts the text present in the form of either image or PDF's into speech (audio). So that visual impaired users can access the content by

listening to relative audio of the content as many times as they want. This system is developed with the intention to make the learning process much easy for visual impaired users.

The main objective of the proposed system is to bridge the gaps the visually impaired students face and to develop an application that reads the textual content and educational data from the given input and converts it into the audio format using Text-to-speech technology.

The basic idea behind the proposed application is helping blind students, also the normal users to understand text better via audio. The objective is to show that student's disability shouldn't be the major reason to hold them from pursuing their education and goals.

LITERATURE SURVEY

In this section, a reference to the relevant past literature that use the various text-to-speech techniques for conversion of mathematical data is mentioned. Most of the researchers concentrate on developing an efficient text-to-speech system that will convert any mathematical data like equations, formulas and problems into speech for blind person.

Sachin Kulkarni and Dr.Debajyoti Mukhopadhyay [1] proposed one text-to-speech tool 'Math Says', that will help legally blind students and also the normal users; understand mathematics better via audio form. The written mathematics material is examined as an image. Then that image is forwarded for the text detection. The detected text is then taken out from that image. The extracted text will be given to the TTS conversion component and as a result, the audio file will be generated which speaks the text in that image. The challenge in this work is how to tackle the formulae which are not in proper English form. For this obstacle, the table will be maintained which has each formula and its corresponding text. Whenever the formula is encountered, the same will be searched in the table and the corresponding text will be extracted and spoken out loud. In this way, this system works for the betterment of understanding the mathematics.

K´evin Vythelingum Yannick Est`eve and Olivier [2] proposed a method to automatically detect grapheme-to-phoneme conversion errors by comparing contrastive phonemisation hypothesis. A lattice-based forced alignment system is implemented, allowing for signal dependent phonemisation. He also implemented a sequence-to sequence neural network model to obtain a context-dependent grapheme-to-phoneme conversion. On a French dataset, it is shown that it can detect to 86.3% of the errors made by a commercial grapheme-to-phoneme system. Moreover, the amount of data annotated as erroneous is kept under 10% of the total evaluation data. The time spent for phoneme manual checking can thus been drastically reduced without decreasing significantly the phonemic transcription quality.

Manjare Anil, S.D. Shribahadurkar and Shaikh Shadab Shakil [3] proposed a new method for a Devanagari (Marathi) Text to Speech system. As there is very less work has been done in Devanagari TTS systems especially for Marathi text. A new method is proposed with mapper and combiner in order to compose the Marathi TTS system. The Marathi TTS system is build using existing English TTS Engine. Marathi input text is the compared with the text present in the database using simple Linear search algorithm then it is provided as input to the Existing English TTS. Contently Concatenative speech synthesis is the method mostly used in TTS systems. The proposed system introduces new concwhich is more feasible and easier than the earlier methods used for Marathi Text To-Speech. Also, the proposed method provides maximum accuracy for text mapping.

Partha Mukherjee and Soumen Santra [4] developed a useful text-to speech synthesizer in the form of a simple application that converts inputted text into synthesized speech and reads out to the user which can then be stored as an mp3 file. They developed a Text-to speech synthesizer is developed that converts text into spoken word, by analyzing and converting it using Natural Language Processing (NLP) and then using Digital Signal Processing (DSP) technology to convert this processed text into synthesized speech representation of the text.

PROPOSED SYSTEM

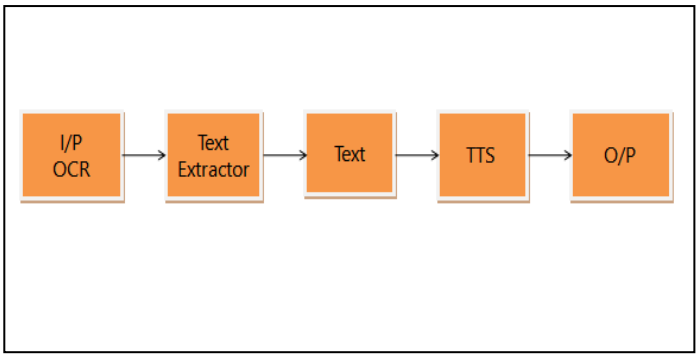

Fig 1. Block Diagram

The primary goal of the system is to convert the text into audio from. The system will take the PDF as an input and use the OCR system to convert the scanned book into text and then the same will be converted to audio form, later storing it in the database. The system is developed in such a way that it can be used by visually impaired students. The challenge was to develop a system making it useful to such users who are not capable to view the system due to sight disorder or poor vision power. The system's GUI is designed in such a way which makes the whole navigation process easy for such category of users. System consists of large tabs and voice guidance on mouse hover on every tab. A textbook surfing interface developed which will text aloud and having facility of moving back and forth using "previous page" functionality as well as "next page" functionality. Also reading desired page by giving page number as input is provided in the system. The system is constructed with fully voice guidance throughout all web pages, making it easier for the user to navigate. The system comprises of a GUI so that the user can provide different input using the mouse hover action to provide input actions such as play, pause, next page, previous page to interact with the system. The interface is designed by focusing mainly on visual impaired user making it very simple, responsive and mobile friendly so it can be used in variety of platforms. The output of the system is the audio conversion of the input text in human voice format.

The TTS is a system for conversion of text documents to audio. The technology helps users to read the written material by listening to the audio output generated by the system. In TTS system the text is taken as input and then converts graphemes to phonemes and finally converting the phoneme to speech. Graphemes are the basic unit of writing system of a language.[1] Phonemes are the smallest meaningful sound element of the language. The TTS system converts text to a phonemic representation and then converts the phonemic representation to waveforms that can be obtained in output as sound. There have been developed many approaches and algorithms to implement TTS conversion.

Text and diagrams are detected using the input, which could be in the form of PDF or IMAGE (JPEG). Reading and writing text is inherently difficult for visually impaired people.

ALGORITHM:

1. Start
2. Input text
3. OCR conversion and segregate input text
4. Store it into the database and assign ID for relative text
5. Parse the text with the suitable ID to convert into speech format
6. Output is in human voice format for the given input text
7. End

RESULTS

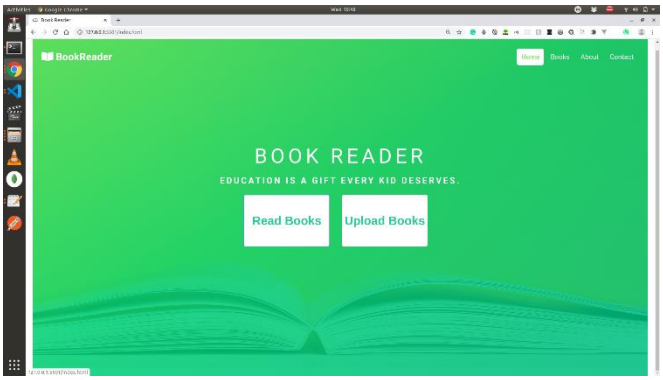

Fig 2. Home Page

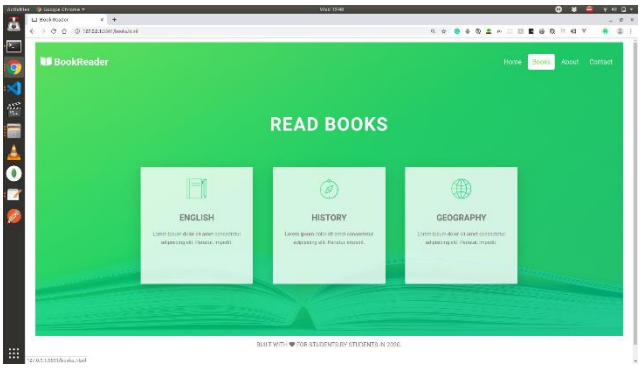

Fig 3. Book Selection Page

Fig 4. English Textbook

SCOPE AND FUTURE WORK

Currently the system only supports English language, but the system can be scaled, and support can be provided for multiple languages, so that it can be used by various medium of instruction used by educational institutes. Therefore, students other than English medium can also use the system for their convenience; hence targeting large number of users. The system maintains a NoSQL database to store PDF's and audio to identify similar format of data, the system can be further extended to support runtime. The application interface is responsive and user friendly to ease the learning in best possible form. The system would be constantly upgraded for new features and regularly tested for errors and bugs, thus providing more accuracy and less error prone environment. NLP capabilities to be included providing the ability of a computer to understand, analyze, manipulate, and potentially generate human language. The TTS technology further can be developed for more accuracy.

CONCLUSION

The system would help the visually impaired to learn and understand whatever they want at anywhere and at any point of time. They do not need any mentor or a guide always to help them study. Thus, the system is

developed with a motivation to increase interest in studies of visually impaired students by making knowledge extraction process simple.

REFERENCES

1) Sachin Kulkarni and Dr. Debajyoti Mukhopadhyay, "TEXT-TO-SPEECH CONVERSION FOR MATHEMATICAL DATA", April 2014
2) KévinVythelingum, Yannick Estève, Olivier Rosee, "Error detection of grapheme-to-phoneme conversion in text-to-speech synthesis using speech signal and lexical context", 2017 IEEE Automatic Speech Recognition and Understanding Workshop (ASRU), December 2017
3) ManjareChandraprabha Anil, S. D. Shirbahadurkar, Shaikh Shadab Shakil, "A mapper and combiner-based Marathi text to speech synthesis using English TTS Engine", 2015 Annual IEEE India Conference (INDICON), December 2015
4) ParthaMukherjee, SoumenSantra, SubhajitBhowmick, Ananya Paul, PubaliChatterjee ArpanDeyasi, "Development of GUI for Text-to-Speech Recognition using Natural Language Processing", 2018 2nd International Conference on Electronics, Materials Engineering &Nanotechnology (IEMENTech), May 2018