

Listen Your Reads

Reena Somani¹ Sumit Janawlekar² Niranjan Jadyar³ Komal Kale⁴ Priyanka Mohite⁵
¹Lecturer ^{2,3,4,5}Student

^{1,2,3,4,5}Department of Information Technology
^{1,2,3,4,5}Atharva College of Engineering, Mumbai, India

Abstract— The Application is built using Vb.Net; Our Application is used to read the multiple documents such as pdf file, word file, text file and rtf file. The Application can be used to view multiple documents, It uses the Text to Speech Engine for Reading the particular document that is being uploaded by the user. The User can Select different speech readers such as male and female, The user has also control over the volume and the rate of the speech function of our application. The User can pause the speaker from reading the given (uploaded) file, the user can select the particular text that he/she want the application to read, When the Speaker Reads the Text from the uploaded file that particular text or word gets highlighted. The application Aims to Read/View Multiple document with some special features such as STOP, PAUSE, VOLUME, RATE OF SPEECH, selection of voice(Male/Female).

Key words: PDF, TTS, .NET, converter

I. INTRODUCTION

A user with low vision wishes to access a multiple document with a large quantity of small text. The user first increases the magnification level of the text, but finds that her eyes are becoming strained. She opts for the application to read her the content of the text file. Application is very much helpful for the blind person or the person who has their eyes vision low.

Text-To-Speech (TTS) synthesizer is a computer based system that should be able to read text from any formatted. In this paper, we have proposed a rule based concatenative synthesis method for generating speech from PDF document. We first discuss PDF to unicode text conversion process and then give an algorithm for the unicode text to speech conversion.[2]

Text-to-speech (TTS) synthesis technology gives machines the ability to convert arbitrary text into audible speech, with the goal of being able to provide textual information to people via voice messages. Key target TTS applications in communications include: voice rendering of text-based messages such as email or fax as part of a unified messaging solution, as well as voice rendering of visual/text information such as web pages. In the more general case, TTS systems provide voice output for all kinds of information stored in databases such as phone numbers, addresses, navigation information, restaurant locations and menus, and movie guides. Ultimately, given an acceptable level of speech quality, TTS could also be used for reading books i.e., Talking Books.[3]

II. LITERATURE REVIEWED

Ajinkya Domale, Bhimsen Padalkar, Raj Parekh, M.A. Joshi presented, Printed text from textbook having English, Marathi or Braille script will be taken as input in the form of an image which will be converted into plain editable text with the help of Optical Character Recognition (OCR). This plain text will be then fed to Text to Speech (TTS) converter which will generate the audio output file in English or Marathi language corresponding to the input text image script[1]

Rafiqul Islam, Ram Shanker Saha, Ashif Rubayat Hossain proposed bangla text to speech conversion. Here our target is to generate almost natural speech from Bangla PDF text. For this we have proposed a speech synthesizer, which performs two major tasks. One is PDF to unicode text conversion and another is unicode text to speech conversion[2].

Azhar Ali Shah, Abdul Wahab Ansari and Lachhman Das proposed, Text to Speech Synthesis along with the Speech Recognition is widely used throughout the world to enhance the accessibility of the information and enable even the disabled persons to interact with the computers in order to get the potential benefit from this high-tech revolution. In this paper we introduce a bi-lingual novel[3].

III. PROPOSED SYSTEM

Web services allow objects to be located anywhere on the Internet and to be called from any application across the Internet (no more trying to get DCOM configured). Of course, extending applications across the Internet will increase security risks. The .NET Framework has many security features built-in to it to protect your applications.

Visual Studio .NET is now built on top of the .NET Framework. The .NET Framework takes application development to viewing the Internet as your new operating system. Your applications will no longer recognize hardware as a boundary limitation. This is an evolution of the Windows DNA model. This new framework is built on open Internet protocols for a standardized interoperability between platforms and programming languages. The .NET Framework will also allow the creation of new types of applications. Applications will now run using the Common Language Runtime (CLR). All .NET applications will use this same runtime environment, which allows your Visual Basic applications to run on equal ground with other languages.

The CLR allows Visual Basic to provide inheritance and free threading, whose absence created two glaring limitations to Visual Basic Applications. Visual Basic .NET is object-oriented. Everything is now an object, and every object is inherited from a standard base class.

IV. SYSTEM IMPLEMENTATION

Applications can control text-to-speech (TTS) using interopspeechlib.dll Component Object Model (COM) interface. Once an application has created an object, the application only needs a call to generate speech output from some text data. In addition, this interface also provides several methods for changing voice and synthesis properties such as speaking rate output volume and changing the current speaking voice.

Special SAPI controls can also be inserted along with the input text to change real-time synthesis properties like voice, pitch, word emphasis, speaking rate and volume. This synthesis markup sapi.xsd, using standard XML format, is a simple but powerful way to customize the TTS speech, independent of the specific engine or voice currently in use.

This method can operate either synchronously (return only when completely finished speaking) or asynchronously (return immediately and speak as a background process).

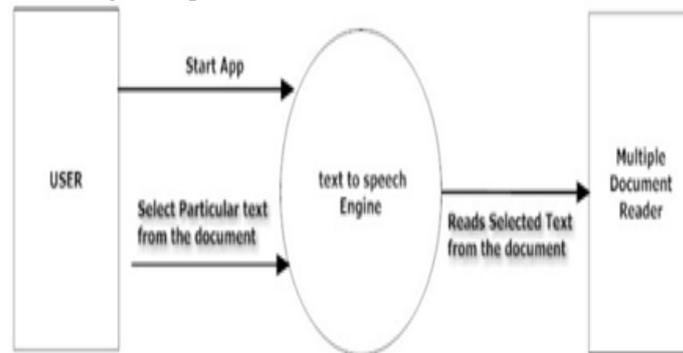


Fig. 1: Block Diagram

Collaboration Diagram

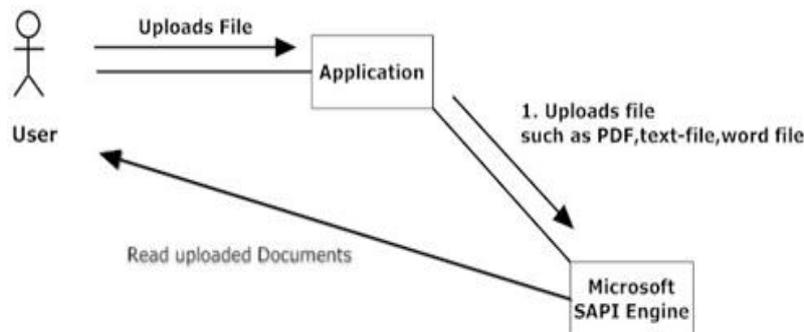


Fig. 3: Collaboration Diagram

V. CONCLUSION

In this paper, we have presented PDF to unicode text Conversion process and proposed an algorithm for converting unicode text to speech. We will also analyze the performance of the proposed algorithm and the existing method in terms of time complexity and thus the proposed algorithm will take less time than that of existing method. Applications will run using the Common Language Runtime (CLR). Such system will greatly help illiterate visually impaired people and people working along with them [1]. This is a PDF to audio desktop converter application. Moreover, this application is a great tool for students who want to study their reading materials by listening to them.

VI. EXPECTED RESULTS

In this paper we present away of English text to speech conversion. Here our target is to generate almost natural speech from English PDF text. For this we have proposed a speech synthesizer, which performs two major tasks. One is PDF to unicode text conversion and another is unicode text to speech conversion. Applications will now run using the Common Language Runtime (CLR). Also the whole design has a simple hardware structure and fast running speed.

ACKNOWLEDGMENT

It gives us great pleasure in presenting this project report titled: "Listen Your Reads." On this momentous occasion, we wish to express our immense gratitude to the range of people who provided invaluable support in the completion of this project. Their guidance and encouragement has helped in making this project a great success. We express our gratitude to our project guide Prof Reena Somani, who provided us with all the guidance and encouragement and making the lab available to us at any time. We also would like to deeply express our sincere gratitude to Project coordinators. We are eager and glad to express our gratitude to the Head of the Information Technology Dept. Prof Neelima Pathak, for her approval of this project. We are also thankful to her for providing us the needed assistance, detailed suggestions and also encouragement to do the project. We would like to deeply express our sincere gratitude to our respected principal Prof. Dr. Shrikant Kallurkar and the management of Atharva College of Engineering for providing such an ideal atmosphere to build up this project with well-equipped library with all the utmost necessary reference materials and up to date IT Laboratories. We are extremely thankful to all staff and the management of the college for providing us all the facilities and resources required.

REFERENCES

- [1] Ajinkya Domale, Bhimsen Padalkar, Raj Parekh, M.A. Joshi College of Engineering, Pune, National Highway 4, Shivaji Nagar, Pune, Maharashtra – 411005
- [2] Md. Rafiqul Islam, Ram Shanker Saha, Ashif Rubayat Hossain Computer Science and Engineering Discipline Khulna University Khulna-9208, Bangladesh
- [3] Az. har Ali Shah, Abdul Wahab Ansari and Lachhman Das Institute of IT, University of Sindh, Jamshoro, Pakistan
- [4] R. M. K. Sinha, "A Syntactic pattern analysis system and its application to Devanagari script recognitio," Ph.D. Thesis, Dept. Elect. Eng., Indian Institute of Technology, Kanpur, India, 1973.
- [5] R. M. K. Sinha and H. Mahabala, "Machine recognition of Devanagari script" IEEE Trans. Syst. Man Cybern., vol. 9, no. 8, pp. 435–441, Aug.1979.
- [6] S. Palit and B. B. Chaudhuri, "A feature-based scheme for the machine recognition of printed
- [7] B. Yegnanarayana, S. Rajendran, V. R. Ramachandran, and A. S. Madhukumar, "Significance of knowledge sources for a text-to-speech system for Indian languages," Sadhana, pp. 147–169, 1994.
- [8] S. K. D. Mandal and B. Pal, "Bengali Text to Speech Synthesis: A Novel Approach for Crossing Literacy Barrier", CSI-YITPA(E), 2002.