Text to speech conversion

Ashwin Sinha – 201IT112
Information Technology
National Institute of Technology Karnataka
Surathkal, India 575025
Email: ashwinsinha.201it112@nitk.edu.in

Vaibhav Kumar Hapani – 201IT124
Information Technology
National Institute of Technology Karnataka
Surathkal, India 575025
Email : vaibhavhapani.201it124@nitk.edu.in

Kunal Singh Tawar – 201IT130
Information Technology
National Institute of Technology Karnataka
Surathkal, India 575025
Email : kunalsinghtawar.201it130@nitk.edu.in

Abstract - The world is moving towards digitization, so are the means of communication. Phone calls, emails, text messages etc. have become an integral part of message conveyance in this tech-savvy world. In order to serve the purpose of effective communication between two parties without hindrances, many applications have come to picture, which acts as a mediator and help in effectively carrying messages in form of text, or speech signals over miles of networks. Most of these applications find the use of functions such as articulatory and acoustic-based speech recognition, conversion from speech signals to text, and from text to synthetic speech signals, language translation amongst various others. In this review paper, we'll be observing different techniques and algorithms that are applied to achieve the mentioned functionalities.

INTRODUCTION

Passing on information, to the right person, and in the right manner is very important. A path-breaking innovation has recently come to play in the SMS technology using the speech recognition technology, where voice messages are being converted to text messages. Quite a few applications used to assist the disabled make use of TTS, STT, and translation. They can also be used for other applications, taking an example: Siri an intelligent automated assistant implemented on an electronic device, to facilitate user interaction with a device, and to help the user more effectively engage with local and/or remote services. Speech synthesis works well in helping convert tokenized words to artificial human speech.

LITERATURE SURVEY

Chaw Su Thu Thu, Theingi Zin, 2014, Implementation of Text to Speech Conversion, International Journal of Engineering Research & Technology (IJERT) Volume 03, Issue 03 (March 2014),

In this work, there were two main parts:
• Optical Character Recognition System for Paper Text
• Text to Speech Conversion

OCR includes these three parts
• Template file Creation
• Creating the Neural Network
• Character Recognition

Template file creation. Letter A to Z and number 0 to 9 images are collected. Each image is changed into 5x7 character representation in single vector by using step 1 to 5 as described in the character recognition section. These data are saved as data file for training in neural network.

Creating the neural network. A feedforward neural network is used to set up for pattern recognition with 25 hidden neurons. After creating the network, the weights and biases of the network are also initialized to be ready for training. The goal is assigned between 0.01 and to 0.05. The created Neural Network is trained by using data file and target file. The neural network has to be trained by adjusting weight and bias of network until the performance reaches to goal.
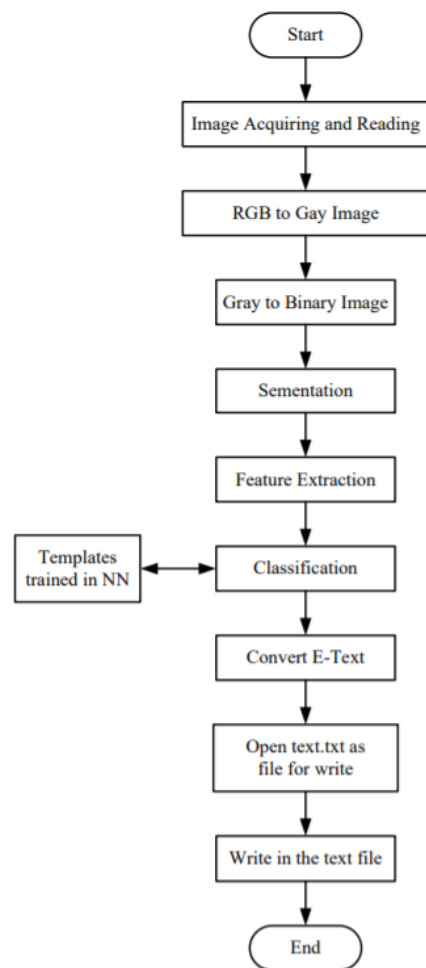

Flow Chart of OCR

The following steps are implemented for character recognition.
• Firstly acquire the character image and the image was read.

• Second step is preprocessing step. In this step firstly the image is converted into gray scale. Then this gray image is converted into black and white image (binary image). Firstly the threshold is counted in gray image then according to that threshold it is converted into black and white image.
• Find the boundary of the character image. Crop the image to the edge.
• Character is extracted and resized in this step. Letters are resized according to templates size.
• The resized binary image is changed into 5 x 7 character representation in single vector.
• Load templates that it can be matched the letters with the templates.
• Open the text.txt as file for write.
• Write in the text file and concatenate the letters.
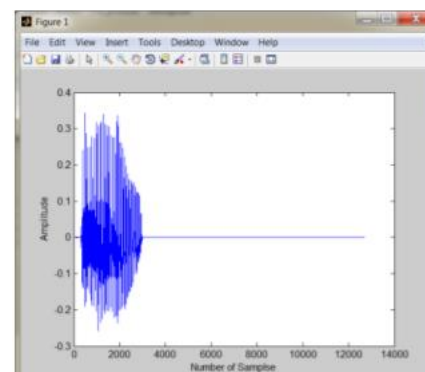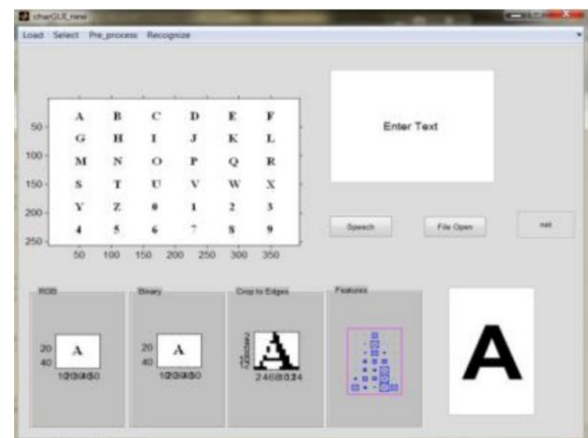
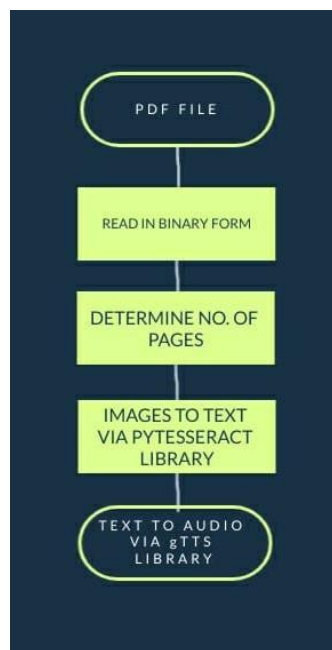Character "A" converted into text "A" sound wave

PROBLEM STATEMENT

To develop an efficient method to convert text into speech

Objective

- To extract text from pdf and convert into speech
- It should recognize both capital as well as small letters.
- It should recognize numbers as well

Methodology

Flowchart of the model

We have implemented Text-To-Speech method for both text file as well as pdf file

The proposed model here first reads the file in binary mode
Then it will determine the total number of pages in the pdf file
A speaker will be initialized with the help of pyttsx3 library.
The default voice of the speaker is male voice but it can be female voice also.

Then our program will read every page one by one and extract the texts present in the current page at each time.
Then all words will be converted from text to speech one by one and will be spoken as a whole string separately.

```python
# Import the Gtts module for text
# to speech conversion
from gtts import gTTS

# import Os module to start the audio file
import os

fh = open("test.txt", "r")
myText = fh.read().replace("\n", " ")

# Language we want to use
language = 'en'

output = gTTS(text=myText, lang=language, slow=False)

output.save("output.mp3")
fh.close()

# Play the converted file
os.system("start output.mp3")
```
Code for a .txt file

```python
# To access voice property of the speaker:
voices = speaker.getProperty('voices')

# Set the speaker's gender: 0-> Male (default), 1-> Female
speaker.setProperty('voice', voices[1].id)

# Iterate through the pages you want to access
# For accessing specific pages: Iterate through the corresponding pageindices
# Note: Index of first page-> 0
# Here, entire PDF is accessed:
for num in range(pages):
    # To read current page index:
    page = pdfReader.getPage(num)
    # To extract the text present in current page:
    text = page.extractText()
    # say() function takes a string as the parameter and then queues the same
    #to be converted from text-to-speech
    speaker.say(text)
    # runAndWait() function blocks the engine instance until all the currently
    #queued commands are processed
    speaker.runAndWait()

# To save the audio output as a MP3 file, within this project:
# Make use of any MP3 player to access this recording whenever required
speaker.save_to_file(text, 'audio.mp3')
```
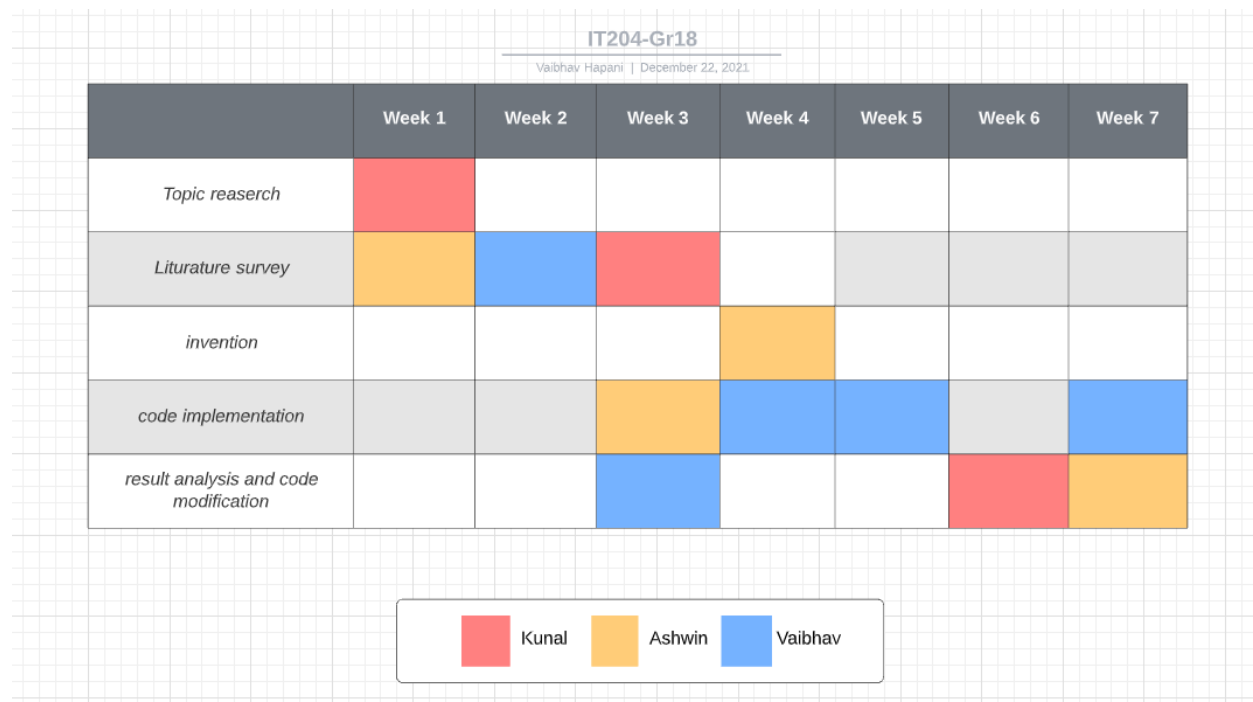Code for a .pdf file

CONCLUSION

This Text-to-Speech model can change the text from pdf or text file to speech successfully. It recognizes alphabets and numbers successfully.

It reads every word separately in other words it sees every single word as a statement.
There is not a specific limit for the length of pdf. Program just reads until it's finished.
It does not recognize the logos or pictures.

If the text involves equations then reader can't properly read the equations.
The method is good for simple text but not for scientific papers.

INDUVIDUAL CONTRIBUTION

IT204-Gr18
Vaibhav Hapani | December 22, 2021

	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7
Topic reaserch	Kunal						
Liturature survey	Ashwin	Vaibhav	Kunal				
invention				Ashwin			
code implementation			Ashwin	Vaibhav	Vaibhav		Vaibhav
result analysis and code modification			Vaibhav			Kunal	Ashwin

Legend: Kunal | Ashwin | Vaibhav

Gantt Chart

IMPLEMENTED/BASE PAPER

Chaw Su Thu Thu, Theingi Zin, 2014, Implementation of Text to Speech Conversion, International Journal of Engineering Research & Technology (IJERT) Volume 03, Issue 03 (March 2014),

References

1. D. Mahanta, B. Sharma, P. Sarmah and S. R. M. Prasanna, "Text to speech synthesis system in Indian English," *2016 IEEE Region 10 Conference (TENCON)*, 2016, pp. 2614-2618, doi: 10.1109/TENCON.2016.7848511.

2. D. S. S. De Zoysa, J. M. Sampath, E. M. P. De Seram, D. M. I. D. Dissanayake, L. Wijerathna and S. Thelijjagoda, "Project Bhashitha - Mobile Based Optical Character Recognition and Text-to-Speech System," *2018 13th International Conference on Computer Science & Education (ICCSE)*, 2018, pp. 1-5, doi: 10.1109/ICCSE.2018.8468858.

3. M. H. O'Malley, "Text-to-speech conversion technology," in *Computer*, vol. 23, no. 8, pp. 17-23, Aug. 1990, doi: 10.1109/2.56867.

4. Ainsworth, W., "A system for converting English text into speech," Audio and Electroacoustics, IEEE Transactions on , vol.21, no.3, pp. 288-290, Jun 1973