

VOIZR- A SOCIAL NETWORK FOR VOICES

¹Ms. S. Pavithra, ²Ms. K. R. Rooshmica, ³Ms. M. Priya, ⁴Ms. R. Kannamma

^{1,2,3}B.E. Students, ⁴Assistant Professor
Department of CSE Engineering, PEC, Chennai, India.

ABSTRACT: The basic idea is that anyone who have this specific application in their mobile can convert articles to speech using Text-to-Speech synthesis(TTS) and read images using OCR algorithm. Our application converts inputted text into synthesized speech and reads out to the user which can then be saved as 3GP(3rd Generation protocol) audio file and listened offline. Image reading is done with Optical Character Recognition (OCR). It is implemented using Tesseract that can be integrated with android. Using this application, the user can either type the text or load an image or copy the contents from URL which can be saved as audio file and used in future. Thus, the proposed application uses Tesseract and Optical Character Recognition (OCR) which helps the user to hear the articles rather than reading during travelling and it also helps the visually challenged people to hear the articles.

Keywords— Text-to-Speech, Optical Character Recognition, Tesseract, Generation Protocol, Uniform Resource Locator .

1. INTRODUCTION

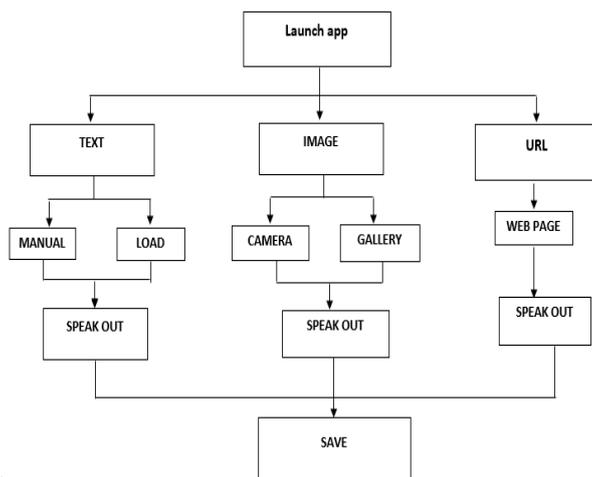
The application field of synthetic speech is growing fast while the quality of TTS systems is also increasing steadily. Speech synthesis systems are becoming more affordable for common customers, which makes these systems more suitable for everyday use. For example, TTS systems may increase employing possibilities for people with communication difficulties. Probably the most important and useful application field in speech synthesis is the reading and communication aids for the blind. Before synthesized speech, specific audio books were used where the content of the book was read into audio tapes. In recent years, OCR (Optical Character Recognition) technology has been applied throughout the entire spectrum of revolutionizing the document management process. OCR has enabled scanned documents to become more than just image files, turning into fully searchable documents with text content that is recognized by computers. With the help of OCR, people no longer need to manually retype important documents while entering them into electronic databases. Instead, OCR extracts relevant information and enters it automatically. The result is accurate, efficient information processing in less time.

2. PROPOSED SYSTEM

The Text-to-Speech synthesis converts the multiple lines into speech using TextToSpeech class imported in android and the converted speech is saved in the device that can be used in future. The user can also load the text files from the device and convert them. Optical character recognition (OCR) algorithm is used to extracting the text from images which can be handwritten, signboards etc. The user can load the images from the device or open the camera to take photos instantly and it can be converted into speech. The audio can also be saved. URL option allows the user to copy the contents from the web page using clipboard and paste them. The pasted text can then be converted to speech and saved in the device.

3. SYSTEM ARCHITECTURE

The system architecture of the application starts with launching the application initially. Once the application is launched the user will be prompted to choose the option of his choice. If the choice is TEXT, the user will be directed to the page where he can manually type the text and save the speech. The user can also load the text files form the device that can be read by TTS. If the choice is IMAGE, then the user has two options to open camera or load the image that has to be read. The OCR recognizes the image and views the content of the image in text format which can be read by TTS. If the choice of the user is URL, then the user has the option to copy the contents of webpage using URL and paste the contents which can be read and converted to speech.



4. MODULES DESCRIPTION

4.1 TEXT TO SPEECH SYNTHESIS

A Text-To-Speech (TTS) synthesizer is a computer-based system that should be able to read any text aloud. Android provides TextToSpeech class for this purpose. In order to use this class, you need to instantiate an object of this class and also specify the initListener. The user can type the article manually and a listen later option can be used to listen audio file later which is in .3gp format. They can listen text from Documents by loading them from the device and save them as audio file. The pitch rate and speed rate can be adjusted according to user's need.

4.2 IMAGE TO SPEECH

Optical character recognition (OCR) is a technology that enables one to extract text out of printed documents, captured images, etc. Android currently doesn't come pre-bundled with libraries for OCR. Tesseract is a well-known open source OCR library that can be integrated with Android apps. we can use tess-two library, which is Tesseract with a Java native interface layer over it, to compile on Android platforms. eng.traineddata is language file that can be created with tool combine_tessdata that is in tesseract directory. An Image to Text Synthesis is conversion of a captured images from our devices default camera or loading images from our device into Text format by the Tesseract. Then the converted article can be saved.

4.3 URL TO SPEECH

On clicking URL button, it displays the web page where the URL can be typed. This class is the basis upon which you can roll your own web browser or simply display some online content within your Activity. It uses the WebKit rendering engine to display web pages and includes methods to navigate forward and backward through a history, zoom in and out, perform text searches and more. The contents of the web page can be copied using clipboard. A Clipboard is a data buffer used for short-term data storage and/or data transfer between documents or applications used by cut, copy and paste operations. We can copy text from additional sources and paste it in the text area which will be converted into audio file and saved it for future use.

5. EXPERIMENTAL RESULTS

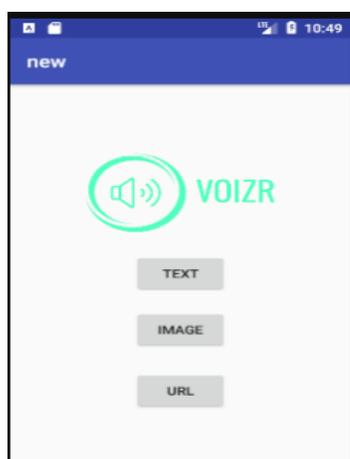


Fig 2: shows the launcher

Once the user launches the app, he can choose from the given options either Text, Image or URL.

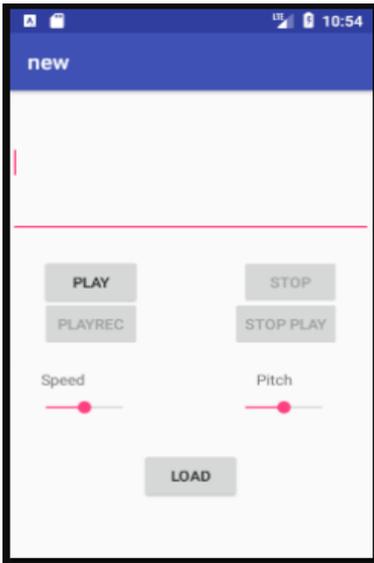


Fig 3: shows text module

When the user clicks text, it directs to the page where he can type the text in text field. On clicking play button, the speech is started. The user can also load the text files from the device.

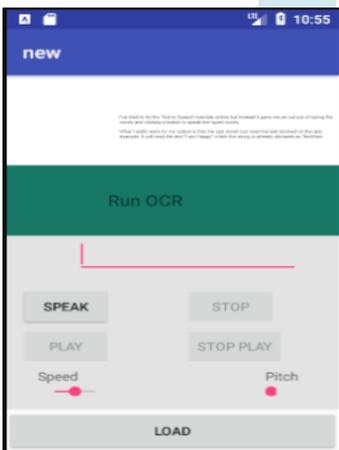


Fig 4: shows Image module

The images can be loaded from the gallery or can be captured by camera instantly. When the run OCR button is clicked, the text in the image is recognized by OCR and are displayed in text field which can later be read.

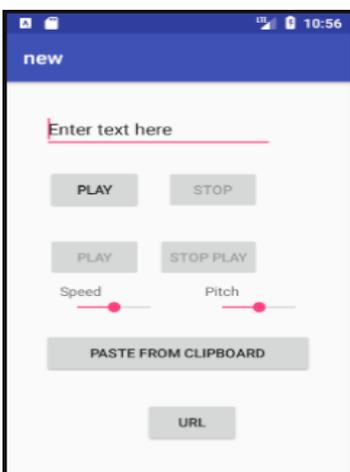


Fig 5: shows URL module

When the URL button is clicked, the page is automatically directed to the webpage, where the user can type the URL and copy the contents from it. Paste from clipboard allows the user to paste the contents to the text field which can be read.

CONCLUSION AND FUTURE WORKS

Thus the proposed application helps the users to increase employing possibilities for people with communication difficulties and probably the most important and useful application field in speech synthesis is the reading and communication aids for the blind. The future work of this application is to detect all kinds of languages with more accuracy of the words pronounced by the user which depends on the language known to the user. The reading application is been good cause for the visually challenged people to know their message by listening. The reading application can detect all languages in future; however text to speech engines for many languages are not available. This will be implemented with the development of text to speech engines. The speech to text provides a good feedback for the user as it is so easy for them to send a message. Proper alignment is essential while holding the device for better results. The future work also includes audio books concept which will be a good business model in the business perspective.

REFERENCES

- [1] Liu Kuei-Chun, Wu Ching-Hung, Tseng, Shau-Yin, Tsai, Yin-Te – Voice Helper: A Mobile Assistive System for Visually Impaired Persons.
- [2] Beg A, Ahmed F, Campbell Hybrid P - OCR Techniques for Cursive Script Languages - Computational Intelligence, Communication Systems and Networks (CICSyN), 2010 Second International Conference on Year: 2010
- [3] Renard Wellnitz Productivity -Text Fairy
- [4] Bissacco A, Cummins M, Netzer Y, Neven H - PhotoOCR- Reading Text in Uncontrolled Conditions - 1-8 Dec. 2013
- [5] Apple Inc. –SIRI, “www.google.com”
- [6] Mande Shen, Hansheng Lei Improving -OCR performance with background image elimination Fuzzy Systems and Knowledge Discovery (FSKD), 2015 12th International Conference on Year: 2015
- [7] Ho J, Ji P, Chen W, Hsieh R - Identifying Google Talk packets Intelligence and Security Informatics, 2009. ISI '09. IEEE International Conference on Year: 2009 8
- [8] Heo G, Yu G, Di L - A reconfigurable open GeoSMS mobile client App design for Android smartphones -Agro-Geoinformatics (Agro-Geoinformatics), 2013 Second International Conference on Year: 2013

