IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661,p-ISSN: 2278-8727, Volume 26, Issue 2, Ser. 1 (Mar. – Apr. 2024), PP 42-50
www.iosrjournals.org

An Enhanced Android-Based Application For Virtually Impaired Individual

Dr. (Mrs.) Ifeyinwa A. Ajah[1] And Okorodudu Joseph[2]
Department Of Computer Science, Ebonyi State University,
Abakaliki Ebonyi State, Nigeria.

Abstract
People who suffer from low vision, sight and visual impairment are not able to see words and letters in ordinary newsprint, books and magazines clearly. This can make the reading process difficult which can disturb the learning process and slow the person's intelligence development. Therefore, an application is needed to help them read. Over the years various applications have been developed to assist the impaired individual to read but still exist some challenges which include speed of sound in the existing application is too fast, and it's difficult to keep up with what the blind is reading, sound or speech is not as natural sounding as a human voice so users find it difficult to comprehend, many different accents and dialects in English, but text-to-speech only has one setting for all voices which makes it not interesting, and no means where various blinds can chat or get any feedback from other people when using text-to-speech (like if they understand what is been said or not). This paper present a new approach where speed of sound can be moderate or slow so that the blind can comprehend easily, there will be natural sounding as a human voice so users can comprehend easily, and finally there will be settings where the blind can select either natural language or setting for all voices to make it interesting. The methodology that was adopted for the development is Object Oriented Analysis and Design Methodology, it was adopted because it is easy to understand and maintain, it provides re-usability, it reduces the development time and cost, and it improves the quality of the system due to program reuse. The application was implemented using android studio, OpenCV and My Structured Query Language (MYSQL) for the database management.

Keywords: *Text-to-speech, human voice, Natural language, Android based, Impaired Individual*

Date of Submission: 01-03-2024 Date of Acceptance: 10-03-2024

I. Introduction

As the use of mobile devices is expanding and affecting various aspects of human life, the number of smartphone users is dramatically increasing. Consequently, the robustness of interaction between smartphone and human is essential for better system performance. By using the speech synthesizer technology for the android, the application presents the modality of text-to-speech responses on android device [1].

A visually impaired person is completely or partially blind. According to an estimate made by the World Health Organization (WHO) 285 million of the population suffer from visual impairment and 39 people were blind and approximately 3% of all the ages in a nation are visually impaired [3]

One of the greatest difficulties faced by a sightless person is the disability to read. Text is present everywhere ranging from bulletin to billboards to digital sections etc. Blind people face a lot of difficulties. There have been developments on mobile phones and computers that assist a blind person by combining computer vision tools with other existing expedient products such as Optical Character Recognition (OCR) system [3]. The proffered system assists blind people by capturing the text and then by reading it to them. Extracting the text present is enacted with OCR. It is a tactic for transformation of images of writings on a label, printed books etc. OCR replaces binary images with texts and also detects white spaces. It also parses the integrity of the recognized text.

As our society farther expands, there have been many supports for second class citizens which are the disabled. One of many supports that are urgent is the guarantee of mobility for blind people. Text-To-Speech has been available for decades (since 1939). Unfortunately, quality of the output-especially in terms of naturalness-has historically been sub-optimal. Terms such as "robotic" have been used to describe synthetic speech. Recently, the overall quality of Text-To-Speech from some vendors has dramatically improved. Quality is now evident not only in the remarkable naturalness of inflection and intonation, but also in the ability to process text such as numbers, abbreviations and addresses in the appropriate context [4].

Text-to-speech (TTS) is a type of speech synthesis application that is used to create a spoken sound version of the text in a computer document, such as a help file or a Web page. TTS can enable the reading of

computer display information for the visually challenged person, or may simply be used to augment the reading of a text message. Current TTS applications include voice-enabled e-mail and spoken prompts in voice response systems [5].

Speech application can be defined as communication between the user and the computer in more natural way or language. As people find speaking naturally is easy, it is the most advantageous to incorporate speech into any natural language processing software. Conversational dialog is a verbal action and it can takes place turn by turn between human and computer [6].

The TTS system gets the text as the input and then a computer algorithm which called TTS engine analyses the text, pre-processes the text and synthesizes the speech with some mathematical models. The TTS engine usually generates sound data in an audio format as the output. The text-to-speech (TTS) synthesis procedure consists of two main phases. The first is text analysis, where the input text is transcribed into a phonetic or some other linguistic representation, and the second one is the generation of speech waveforms, where the output is produced from this phonetic and prosodic information. These two phases are usually called high and low-level synthesis [7].

The input text might be for example data from a word processor, standard ASCII from e-mail, a mobile text-message, or scanned text from a newspaper. The character string is then pre-processed and analyzed into phonetic representation which is usually a string of phonemes with some additional information for correct intonation, duration, and stress. Speech sound is finally generated with the low-level synthesizer by the information from high-level one. The artificial production of speech-like sounds has a long history, with documented mechanical attempts dating to the eighteenth century.

Speech synthesis can be described as artificial production of human speech [8]. A computer system used for this purpose is called a speech synthesizer, and can be implemented in software or hardware. A text-to-speech (TTS) system converts normal language text into speech [9]

Synthesized speech can be created by concatenating pieces of recorded speech that are stored in a database. Systems differ in the size of the stored speech units; a system that stores phones or diphones provides the largest output range, but may lack clarity. For specific usage domains, the storage of entire words or sentences allows for high-quality output. Alternatively, a synthesizer can incorporate a model of the vocal tract and other human voice characteristics to create a completely "synthetic" voice output [10]. The quality of a speech synthesizer is judged by its similarity to the human voice and by its ability to be understood. An intelligible text-to-speech program allows people with visual impairments or reading disabilities to listen to written works on a home computer.

A text-to-speech system (or "engine") is composed of two parts: a front-end and a back-end. The front-end has two major tasks. First, it converts raw text containing symbols like numbers and abbreviations into the equivalent of written-out words. This process is often called text normalization, preprocessing, or tokenization. The front-end then assigns phonetic transcriptions to each word, and divides and marks the text into prosodic units, like phrases, clauses, and sentences. The process of assigning phonetic transcriptions to words is called text-to-phoneme or grapheme-to-phoneme conversion. Phonetic transcriptions and prosody information together make up the symbolic linguistic representation that is output by the front-end. The back-end often referred to as the synthesizer then converts the symbolic linguistic representation into sound. In certain systems, this part includes the computation of the target prosody (pitch contour, phoneme durations), which is then imposed on the output speech [11].

There are different ways to perform speech synthesis. The choice depends on the task they are used for, but the most widely used method is concatentive synthesis, because it generally produces the most natural-sounding synthesized speech. Concatenate synthesis is based on the concatenation (or stringing together) of segments of recorded speech.

Statement of the Problem

Over the years different applications have been developed to help the blind or low vision read but need improvement which include the following:

1. The speed of sound in the existing application is too fast, and it's difficult to keep up with what the blind is reading.
2. The sound or speech is not as natural sounding as a human voice so users find it difficult to comprehend
3. There are many different accents and dialects in English, but text-to-speech only has one setting for all voices which makes it not interesting.

Research Aim/Specific Objectives

The aim of this paper is to introduce an android based for virtually impaired individuals that will make different sound based on typing latters.

The objectives are to:

1. Develop a platform that make typing and reading speed comprehensive
2. Create a platform that Implement natural human tone and the speech for improved comprehension of the speech by the user
3. Create a module that provide a means where users can control the speed of the speech
4. Detect and pronounce a letter and word key selection during typing and create a module where there will be means to select different accents in English

Significance of the Study

This paper will be a useful tool in the hands of the vision impaired individuals; therefore, this goes a long way by creating a text to speech synthesis application. With this android application the blind will be able to know exactly what they are typing and also will be able to read any text by typing the text while the application automatically convert to speech so they will know exactly what they are reading.

Scope of the Study

This paper focus on converting text into spoken word, by analyzing and processing the text 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.

Review of Related Literature

[12], states that work has been done for visually impaired and blind peoples for reading text and detecting objects in their surroundings. In recent years, technology grows very much but this technology is no use for disabled people which provide help. As we know, education is very important in factor in human's life. The education receives normal people that level of education can't receive blind or visually impaired peoples. Blind peoples read books that scripted in Braille language, but the cost of the books is around 1300 rupees. So, imagine that how costs the education for them. Braille language is not efficient as it takes long time to learn. For that assistance is developed for blind people.

The Reader (Blind people) is like a live tutor. Computer vision with IOT raspberry pi is used. By this reader can read whatever they want Raspberry pi acts as a microcontroller for processing. Detection is also done by tensor flow, with the help of camera; images are captured and then analyzed through tensor flow and with various datasets. And with the help of audio output people knows what exactly is around them.

[13], Authors have developed a distance schooling portal for the blind the place they can get entry to a wide variety of training applications alongside with their contents and additionally the length of the path coverage. They have used the modern day technological know-how such as PHP, MySQL for database and JAWS as display analyzing equipment with HTML5 and CSS3 to diagram where the blind can use the device and engage for academic records.

[14], authors have put forth but some other viewpoint of lookup in current-day the place visually impaired schooling and laptop imaginative and prescient to go hand in hand such that it aids the blind to lift out their duties greater easily. They aimed to acquire this by means of creating an utility that can assist young people between the a long time of 6 to 14 years to discover objects besides any different human support. They have used the modern-day technological know-how of artificial talent and pc imaginative and prescient alongside with Deep Region Convolutional Networks alongside with Recurrent Neural Networks in addition to Speech fashions.

[15], there is an increase in the number of visually disabled and blind people, as we all know. While they have many applications and different features, most of them are for the sighted and are not available to blind people. The main problem faced by them is identification of things used in their daily life or to read a text or a book. Therefore, this application provides various functions such as light detection, color detection, object budget recognition, and bank note recognition. Here the color detection was related to our topic; in it color detection approach works on images taken by color detection activity and by using open CV library. The RGB color of the area on the screen that is reached by the user and the color name is recorded to the user by the text to the speech engine available on the Smartphone according to the unique RGB values that are detected. New colors are detected and announced by audible means to the user as the user moves his / her finger on the device. Since the application targets visually impaired persons, the performance results are read out loud.

It uses the cell phone's built-in voice engine to read out the results loudly, so that the user can hear them clearly.

[16], Authors have additionally developed android software which presents a Chabot for the blind so that they can acquire solutions to queries associated to education. To get the solutions, they will be educated for the use of the functions the usage of Google voice search which additionally offers them with answers. These can be both bill layout or can be bought in textual content layout too.

[17] globally, the technologies used to help people with disabilities are either difficult or inexpensive. However current advances in the medical world claim that they provide alleviation and make their lives unbiased however the lack of information coaching and realistically complexity often leads to consumer frustration, this uses the Internet of Things to help the visually impaired (blind) people and to boost their self-confidence not only in a well-known environment, but also in unfamiliar places that would be easily accessible to them. In this preliminary strategy a variety of probabilities for designing a multi-sensory community primarily based on inter objects have been proposed. This method now not solely solves the technical issues associated with handy applied sciences but additionally gives low price and person pleasant functions that can be without problems used to grant comfort.

[18] basically blind people faces a number of challenges while operating smart phones, laptops, etc.

Here there is interface for blind people to perform various activities like placing and receiving messages, emails, etc. So, there are certain requirements user has to meet while developing useable interfaces for blind people. Here one of the most important factor of usability requirements is the division of touch screen in such a way that blind people to locate objects and non-visual items on the screen easily and comfortably. So, this includes dividing touch screen into two, three, four, five, eight, and nine partitions with respect to finger placement. The main aim is to design useable user interfaces for blind people to perform similar activities on Smartphone. The Smartphone services provide many opportunities to blind people to operate Smartphone using screen-reading software and also text to speech system and visual information are transformed into non- visual representation through speech and voice based representation.

[19] "Software and hardware solutions for using the keyboards by blind people", both software and hardware solutions are given for blind users. But mainly focus on software solution. The blind users received information of computer and communication technique by hearing and by touching. So here few software applications which help blind users to communicate with computer. 1. Voice synthesis, it is more efficient, here we can received more information in shorter time. 2. Screen reader, it is an application which help to identify the properties of an object and by using graphical interface; it converts them into text and can be transmitted by a speech synthesis program.3. Convert documents to audio file, by using TTS (Text-To-Speech) we can read some typed text or some programs present in the window, through voice synthesis. Eventually, the app can save the result of conversion into one of the audio formats for later listening. So there are two software solutions discussed in this paper, one of this is Keyboard learning software application: the main aim of this application to create sound when the user presses any key. So whenever we press any key, a character is displayed. So the application compares this character with English database and generates a specific sound. At the same time, the Braille shape of the character is displayed, so after starting this application, the last character in the text box is read, Braille code is searched, and character sound is generated. Here we have one drawback with this application that is; it does not recognize the punctuation marks and the capital letters. And the second solution is Application for testing the typing capabilities of blind person; the objective of this application is to test the rapidity and fairness to writing a text heard by the user. Here we need a person's help to insert a text snipped into that particular box. Here a test is taken in which the spelling errors are calculated, in which all characters are cross-checked and if any character is missing or is extra then all the characters starting from that point are considered wrong. So from these tests, many errors are faced by the blind people. So, this software application needs to be more advanced for the blind users to become useful in typing contests.

[17], the number of blind or visually impaired peoples is so high due to this it effects on the growth and economy of the country. Project works in two ways: 1) Navigation 2) Face Detection. This paper states that it uses GPS module for navigation, Image processing for face detection and Raspberry pi for inputs. GPS and raspberry pi trace the location by latitude and longitude values received by the antenna. OpenCv used for processing of captured images which approximately 20 different images take before processing. LDR used for detection of light and dark but it is inefficient. IR sensor detects the objects by using motion sensor which detects the heat released by our body. After that Raspberry Pi code, the output in voice message. This device helps to give directions for the stored path of destination. It also monitors the GPS current locations.

[15], RFID tags are attached to various objects inside the house for tracking and identification purposes. The user has an RFID reader that can detect the tags in the user's proximity and inform the user about the name of the object on which the detected tag is placed. The user's device then sends the tag-id to the nearby computer. The computer looks up the tag ID in the database that is maintained. This database has an audio clip, which is associated with each tag id, contains the name of the object.

Analysis of the Existing System

Speech synthesis can be described as artificial production of human speech. A computer system used for this purpose is called a speech synthesizer, and can be implemented in software or hardware. A text-to-speech (TTS) system converts normal language text into speech. Synthesized speech can be created by concatenating pieces of recorded speech that are stored in a database. Systems differ in the size of the stored

speech units; a system that stores phones or diphones provides the largest output range, but may lack clarity. For specific usage domains, the storage of entire words or sentences allows for high-quality output. Alternatively, a synthesizer can incorporate a model of the vocal tract and other human voice characteristics to create a completely "synthetic" voice output. The quality of a speech synthesizer is judged by its similarity to the human voice and by its ability to be understood. An intelligible text-to-speech program allows people with visual impairments or reading disabilities to listen to written works on a home computer.

A text-to-speech system (or "engine") is composed of two parts: a front-end and a back-end. The front-end has two major tasks. First, it converts raw text containing symbols like numbers and abbreviations into the equivalent of written-out words. This process is often called text normalization, preprocessing, or tokenization. The front-end then assigns phonetic transcriptions to each word, and divides and marks the text into prosodic units, like phrases, clauses, and sentences. The process of assigning phonetic transcriptions to words is called text-to-phoneme or grapheme-to-phoneme conversion. Phonetic transcriptions and prosody information together make up the symbolic linguistic representation that is output by the front-end. The back-end often referred to as the synthesizer then converts the symbolic linguistic representation into sound. In certain systems, this part includes the computation of the target prosody (pitch contour,

There are different ways to perform speech synthesis. The choice depends on the task they are used for, but the most widely used method is Concatentive Synthesis, because it generally produces the most natural-sounding synthesized speech. Concatenative synthesis is based on the concatenation (or stringing together) of segments of recorded speech. There are three major sub-types of concatenative synthesis.

Analysis of the Proposed System

The proposed system is called the TextToSpeech Robot, a simple application with the text to speech functionality. The system will be developed using android studio. Android studio is used because it's robust and independent platform. The application is divided into two main modules, the main application module which includes the basic GUI components which handles the basic operations of the application such as input of parameters for conversion either via file or direct keyboard input or the browser. This would make use of the open source API called SWT and DJNativeSwing. The second module, the main conversion engine which integrated into the main module is for the acceptance of data hence the conversion. This would implement the API called freeTTS.

TextToSpeech Robot (TTSR) converts text to speech either by typing the text into the text field provided or by coping from an external document in the local machine and then pasting it in the text field provided in the application. It also provides a functionality that allows the user browse the World Wide Web (www) on the application. TextToSpeech Robot is capable of reading any portion of the web page the user browses. This can be achieved by the user highlighting the portion he wants to be read out loud by the TTSR and then clicking on the "Play" button.

TTSR contains an exceptional function that gives the user the choice of saving its already converted text to any part of the local machine in an audio format; this allows the user to copy the audio format to any of his/her audio devices.

The default view for this application being created is the web browser view. This is what shows after the TextToSpeech Robot has loaded. The web browser displays that there is no internet connection on the local machine and so it displays "The page cannot be displayed". Any part of the web browser in the application that is highlighted can be read out loud by the TTSR. The application allows the user to highlight any part of the web page for conversion. The proposed system uses natural human tone and the speech for improved comprehension of the speech by the user and also the application is able to correct any wrong spellings using the concept of Microsoft word, it will display the wrong text in red before correcting it.

Use case Diagram of the Proposed System

The use case diagram in figure 1 explains the functionalities of various users in the system. Here users input text to the app and click on play while the application convert the text to speech and read out the text, users can stop or pause the speech at any time and can also select the ascents of the spoken text and as well reduce the speech volume. The administrator then manages speech and manages dialect.

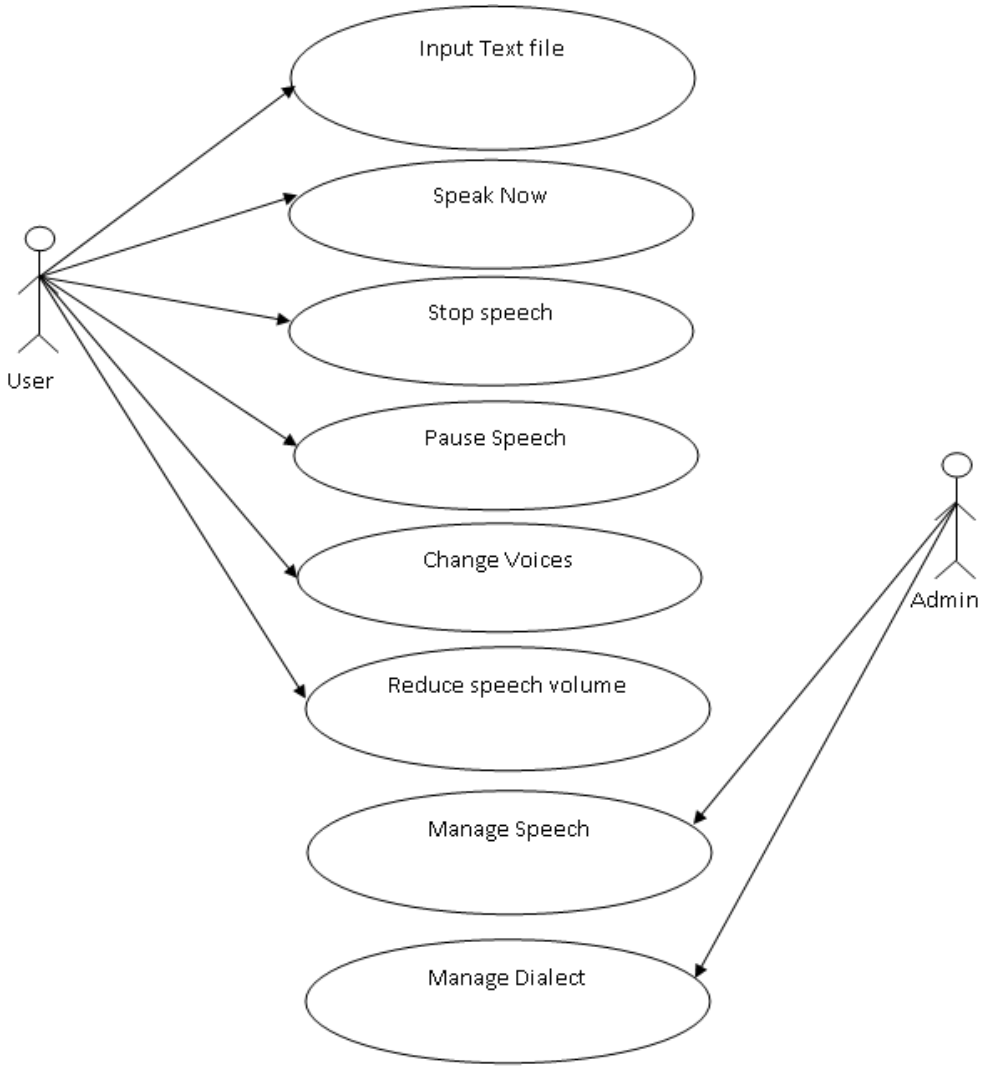

Figure 1: Use case Diagram of the Proposed System

The use case diagram in figure 1 explains the functionalities of various users in the system. Here users input text to the app and click on play while the application convert the text to speech and read out the text, users can stop or pause the speech at any time and can also select the ascents of the spoken text and as well reduce the speech volume. The administrator then manages speech and manages dialect.

Sequence Diagram of the Proposed System

Figure 2 explains the sequence diagram of the proposed system; here the user will install the application input the text while the application convert the text to speech and read it out to the user and user can either pause or stop the speech.

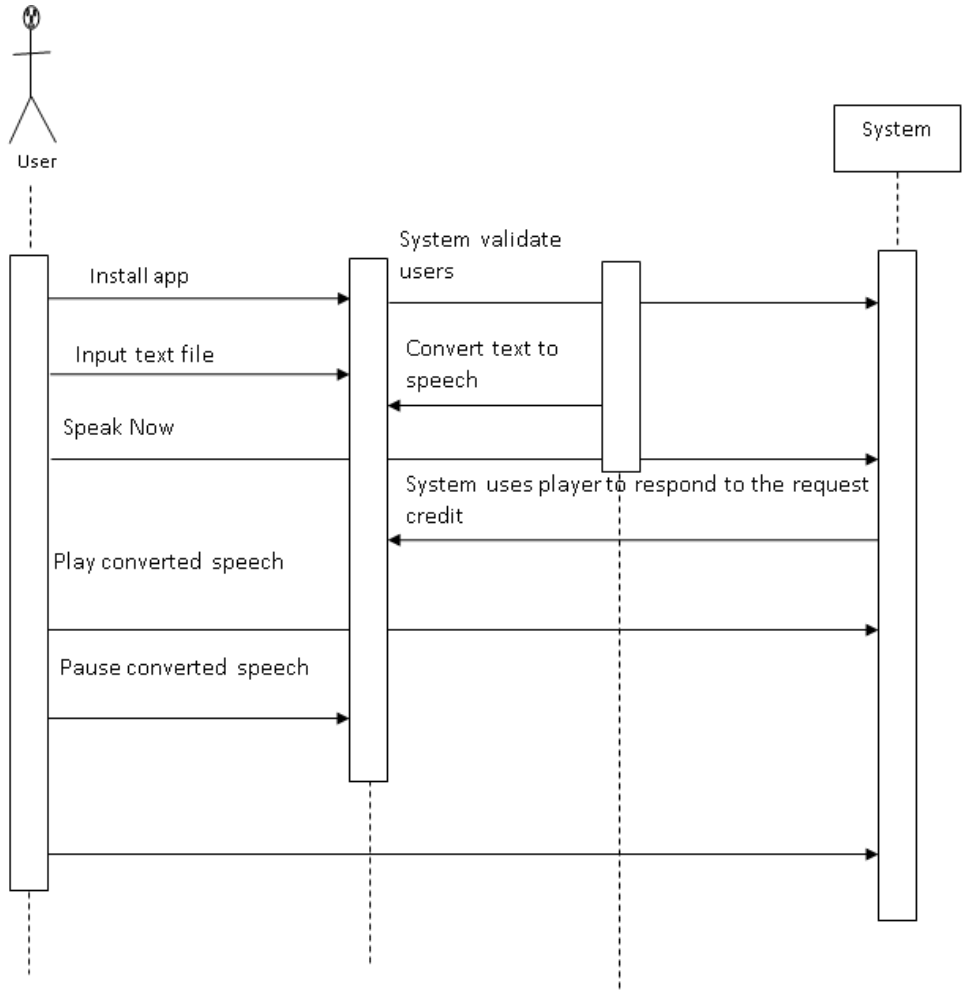

Figure 2: Sequence Diagram of the Proposed System

Figure 2 explains the sequence diagram of the proposed system; here the user will install the application input the text while the application convert the text to speech and read it out to the user and user can either pause or stop the speech.

Design Methodology

Object-Oriented Analysis and Design Methodology (OOADM) will be used to define and document the class hierarchy from which all the system objects are created and object interactions are defined. The use of this methodology will help to produce a better quality software product, in terms of documentation standards, acceptability to the user, maintainability and consistency of software.

Implementation Phase

The user interface implementation is the first interface that displays when a user launch the application, this is the interface where users interact with. It is on the interface that the user navigates to other menus of the application. The interface implementation is shown figure 3:

Figure 3: User Interface Implementation

Figure 3 is the user interface implementation, this is the first interface of the application, and this is the interface that allows users to navigate from one menu to another.

Input Interface Implementation Phase

There is only one method of input to the app which is where a user will enter or type text that will be converted to speech is shown in figure 4.

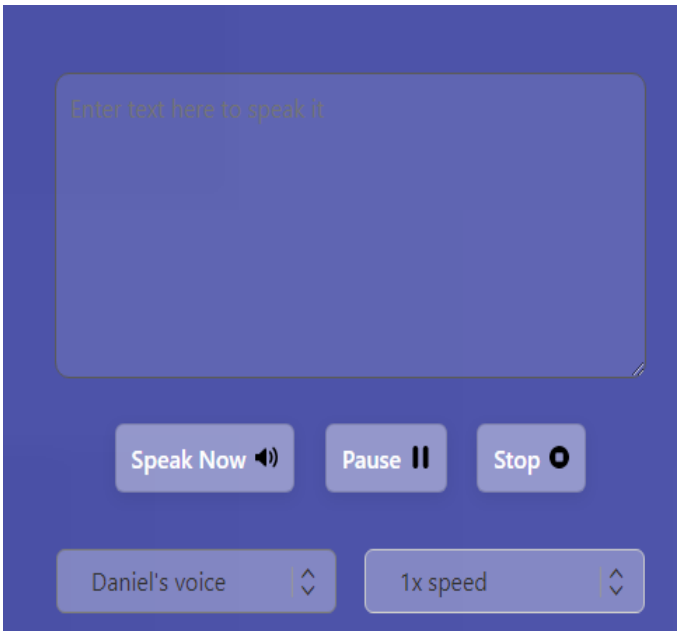

Figure 4: Input Interface Implementation

Figure 4 is the interface that enables the user to input test which the application can read out loud.

Output Implementation Phase

The output implementation is the result from the input the user supplied to the system in other to display the result after the input has been made. After the user must have keyed in the test to the textbox, the

output of the application is that the application will automatically read out the text in form of voice with different intonations.

System Requirement

System requirement specifies the type of hardware and software that will enable the system to run effectively. The application requires an android phone with any version. Its characteristics of mobility, autonomy, collaborative behavior and adaptability makes it use lest space in its environment. The requirements are listed below:

i. System Type: Android based Mobile Phone
ii. Processor: Quad-core 1.7GHz cortex-A53 and quard-core 1.0GHz cortex-A53
iii. Internal memory (RAM): 16/32GB, 2GB RAM
iv. Total virtual Memory: 64MB

The system will run on Android platform. The android platform provides support for user interface management platform using Extended Markup Language (XML); this will provide the user friendly interface for the system and SQlite database server will run on the background of every installed phone.

II. Conclusion

In conclusion all the stated objectives were achieved in which the developed system enable users to enter text while the application read out the text. This thesis made a clear and simple overview of working of text to speech system (TTS) in step by step process. There are many text to speech systems (TTS) available in the market and also much improvisation is going on in the research area to make the speech more effective, natural with stress and emotions. This paper introduces a means where a user can select his or her desired speech and volume before listening.

References

[1]. Archana Balyan, S.S. Agrwal And Amita Dev, (2013) Speech Synthesis: Review, Ijert, Issn 2278-0181 Vol. 2 (2013) P. 57 – 75.
[2]. Awadand H., And Haddad J., (2018),"Intelligent Eye: A Mobile Application For Assisting Blind People", Menacomm-Ieee, 2018, P 6
[3]. Balasuriya, B. K., N. P. Lokuhettiarachchi, A. R. M. D. N. Ranasinghe, K. D. C. Shiwantha, And C. Jayawardena, (2017) "Learning Platform For Visually Impaired Children Through Artificial Intelligence And Computer Vision.", 2017 11th International Conference On Software, Knowledge, Information Management And Applications (Skima), Pp 17, Ieee, 2017.
[4]. Grupta .W, Chi-Chun Hsia, Te-Hsien Liu, And Jhing-Fa Wang, (2006) Voice Conversion Using Duration-Embedded Bi-Hmms For Expressive Speech Synthesis, Ieee Trans. Speech Audio Process, Vol. 14, (2006) Pp. 1109 – 1116.
[5]. Itomoluwa A., And Chandra E., (2017) Speech Processing –An Overview, Int. J. Of Engg. Sci. And Tech., Vol. 4, (2012) P. 2853-2860.
[6]. Jithendra Vepa And Simon King, (2006) Subjective Evaluation Of Join Cost And Smoothing Methods For Unit Selection Speech Synthesis, Ieee Trans. Speech Audio Process, Vol. 14, (2006) Pp. 1763 – 1771.
[7]. Kumar M., And Naveen K., (2016) Pc Linga Chandar, A. Venkatesh Prasad, And K. Sumangali, "Android Based Educational Chatbot For Visually \ Impaired People", 2016 Ieee International Conference On Computational Intelligence And Computing Research (Iccic), Pp. 1-4, Ieee, 2016
[8]. Mark T., And Katherine M., (2015), Developments In Speech Synthesis (John Wiley & Sons, Ltd. Isbn: 0-470-85538-X, 2005).
[9]. Murad, M., Rehman, A., Shah, A.A., Ullah, S., Fahad, M., Yahya, K.M. (2011) "Rfaide—An Rfid Based Navigation And Object Recognition Assistant For Visually Impaired People" In: 7th Inter- National Conference On Emerging Technologies (Icet), Islamabad, Pakistan, Pp. 1– 4 (2011)
[10]. Murali M., And Sharma S., (2020) "Reader And Object Detector For Blind", Ieee, 2020, P 4+
[11]. Nawandar And Gohokar V., (2018) "Design And Development Of Multisensory Smart Assistive Technology For Blind Persons", Ieee, 2018, P 4
[12]. Ning Ma., Rafik A., Goubran, (2006) Speech Enhancement Using A Masking Threshold Constrained Kalman Filter And Its Heuristic Implementations, Ieee Trans. Speech Audio Process, Vol. 14, (2006) Pp.19-32.
[13]. Onaolap J.O., Idachaba F.E., Badejo J., Odu T., And Adu O.I., (2016) In Proc. Of The World Congress On Engineering, (London, Uk. 2014).
[14]. Pande D.D., Praveen M., Kumar, (2017) A Smart Device For People With Disabilities Using Arm7, Ijert, Issn 2278-0181 Vol.3(2014) P. 614 – 618.
[15]. Rohit Deo And Pallavi Deshpande, (2014) Neutral To Emotional Speech Conversion By Pitch Counter Modification For Marathi, Ijert, Issn 2278-0181 Vol. 3 (2014) Pp. 2228-2231.
[16]. Schusro M., And Niazi K., (2019), A Brief History Of Synthetic Speech, J. Speech Communication, Vol. 13, (1993) P. 231-237.
[17]. Yurtay, Nilufer, Yuksel Yurtay, And Fatih M., (2015) "An Education Portal For Visually Impaired.", 171 (2015): 1097-1105
[18]. Zhen-Hua Ling, Korin Richmond, Junichi Yamagishi, And Ren-Hua Wang, (2009) Integrating Articulatory Features Into Hmm-Based Parametric Speech Synthesis, Ieee Trans. Speech Audio Process, Vol. 17, (2009) Pp. 1171 – 1185.