

Mobile Base Sinhala Book Reader for Visually Impaired Individuals

Project ID: TMP-23-198

Project Proposal Report

Jayathunga T.M. - IT20146238

Supervisor: Prof. Koliya Pulasingha

B.Sc. (Hons) Degree in Information Technology Specialization in Information Technology

Department of Information Technology

Sri Lanka Institute of Information Technology

Sri Lanka

March 2023

Mobile Base Sinhala Book Reader for Visually Impaired Individuals

Project ID: TMP-23-198

Project Proposal Report

B.Sc. (Hons) Degree in Information Technology Specialization in Information Technology

Department of Information Technology

Sri Lanka Institute of Information Technology

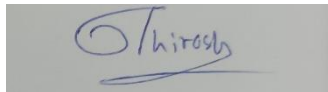
Sri Lanka

March 2023

Declaration of The Candidate & Supervisor

I declare at this moment, that the proposal I am presenting is entirely my work, and I have not incorporated, without proper acknowledgment, any material previously submitted for a degree or diploma at any other university or institute of higher learning. This proposal does not contain any material previously published or written by another person, except where the appropriate acknowledgment has been made in the text.

I am aware of the potential consequences of academic dishonesty and plagiarism, and I am committed to upholding the principles of honesty and intellectual integrity in all my academic endeavors.

Student IT Number	Student Name	Signature
IT20146238	Jayathunga T.M.	

The Supervisor should certify the proposal report with the following declaration.

The candidates mentioned above are currently conducting research for their undergraduate dissertation, under my supervision. As their supervisor, I certify this proposal report.

Signature of the Supervisor

Date

.....

.....

Abstraction

Nowadays, with the advancement of technology, every age group is keenly interested in education. Currently, every sector of the world is developing with Artificial Intelligence (AI) at an unimaginable speed. With this, education is not limited to books, it is being developed with technology to make it easier to learn. Reading books can be the best way for a person to acquire the knowledge he needs. However, visually impaired people can read Braille and nowadays it is possible to do it through Computerized Braille systems or Text to Speech (TTS) systems. Although these types of technologies make things available in English easily accessible to visually impaired people, the equipment available in Sinhala is limited. Furthermore, such devices are very expensive in Sri Lanka, and it is not easy for a visually impaired person to buy these devices and they must spend a lot of money for them. Since Android is the most popular operating system in Sri Lanka, we decided that it would be better to make a mobile application based on it.

Printed books and other documents in the Sinhala language available in Lanka are not improved for reading by visually impaired persons. The main objective of this project is to enable them to have easy access to books and documents. Here, a visually impaired person navigates the app as needed by voice navigation, and the application gives the necessary commands to the visually impaired person. Here the mobile camera performs the necessary commands to accurately capture all the characters on a page of a Sinhala book. After that, optical character recognition separates and identifies the characters from the image with the corresponding characters. It then makes the respective visually impaired person hear clearly through the Festival Synthesizing Framework using Text to Speech. Here he can increase or decrease the speed of reading the book according to the commands. We intended to develop this Mobile Base Sinhala Book Reader to make it easier for visually impaired people by using audio recommendations and Sinhala voice notifications. We hope this will further increase the enjoyment of reading Sinhala books among the visually impaired community living in Sri Lanka.

Key Words: Text to Speech Sinhala, Natural Language Processing, Visually Impaired Individuals, Festival Synthesizing Framework, Speech Synthesis

Table Of Content

Declaration of The Candidate & Supervisor	3
Abstraction	4
List of Tables	6
List of Figures	6
List of Abbreviations.....	6
1. Introduction	7
1.1. Background	7
1.2. Literature Survey	8
2. Research Gap	9
3. Research Problem.....	10
4. Objectives	11
4.1. Main Objective	11
4.2. Sub Objective	11
4.2.1. Text-to-Speech (TTS).....	11
5. Methodology	13
5.1. Overall System Diagram	13
5.2. Individual System Diagram.....	15
6. Project Requirements.....	16
6.1. Functional Requirements	16
6.2. Non-Functional Requirements	16
7. References	17

List of Tables

<i>Table 1 : Comparison Between Existing Systems.....</i>	<i>9</i>
---	----------

List of Figures

<i>Figure 1 : Android and IOS OS Market in Sri Lanka</i>	<i>8</i>
<i>Figure 2 : General Functional Diagram for Text-to-Speech (TTS) System [7].....</i>	<i>12</i>
<i>Figure 3 : Project Overall System Diagram</i>	<i>13</i>
<i>Figure 4 : Text-to-Speech (TTS) Synthesizer.....</i>	<i>15</i>

List of Abbreviations

TTS – Text-to-Speech

OCR – Optical Character Recognition

NLP – Natural Language Processing

WHO - World Health Organization

1. Introduction

1.1. Background

Vision impairments are widespread, and far too often they still go untreated. Globally, at least 2.2 billion people have a vision impairment, and of these, at least 1 billion people have a vision impairment that could have been prevented or is yet to be addressed [1].

Knowledge has emerged as the essential component for effective living in the 21st century. Reading becomes the primary way to learn new things. People who are blind or visually impaired are interested in learning further, but the issue is that they cannot locate all sources of information in braille format because many sources of information are not available in this format. Many technological solutions have been developed in response to this, including text-to-speech and computer-aided braille systems.

Many technological solutions have been developed in response to this, including text-to-speech and computer-assisted braille systems. Most of these systems are widely available in developed countries and these systems are of high price [2]. Additionally, most text-to-speech tools are for the English tongue. However, these amenities are not available to Sri Lankans whose native language is Sinhala and whose level of English literacy is poor. As a result, we require such a mechanism.

Mobile-based technologies are quickly getting acceptance across the globe in the era of modern technological advancements. Even though Sri Lanka is still a developing country, most of its citizens own a smartphone [2]. The mobile phone business is mainly dominated by Android devices, as the worldwide trend also demonstrates. Therefore, the most effective strategy to handle the problems encountered by visually disabled people in Sri Lanka would be a smartphone solution built on Android.

In conclusion, Android-based mobile book readers offer several benefits to Sri Lankans who are blind. The open-source platform's adaptability enables programmers to build specialized apps with support for the Sinhala language and accessibility capabilities. Making mobile book readers for Android is a common option due to the platform's popularity and the accessibility of developer tools. Android-based mobile book readers can reach a larger population of people who are blind because Sri Lanka has a high percentage of Android users.

1.2. Literature Survey

The use of mobile devices has grown considerably in recent years, giving people a more convenient method to obtain information. For those who are blind, there is an increasing demand for portable text readers in Sri Lanka. Due to the open-source nature of the Android platform and the abundance of developer tools, Android-based apps have grown in popularity. This literature review seeks to shed light on the benefits of developing Sinhala-based mobile book readers for visually impaired individuals in Sri Lanka using Android-based mobile applications.

Mobile Operating System Market Share Sri Lanka

Oct 2011 - Feb 2023

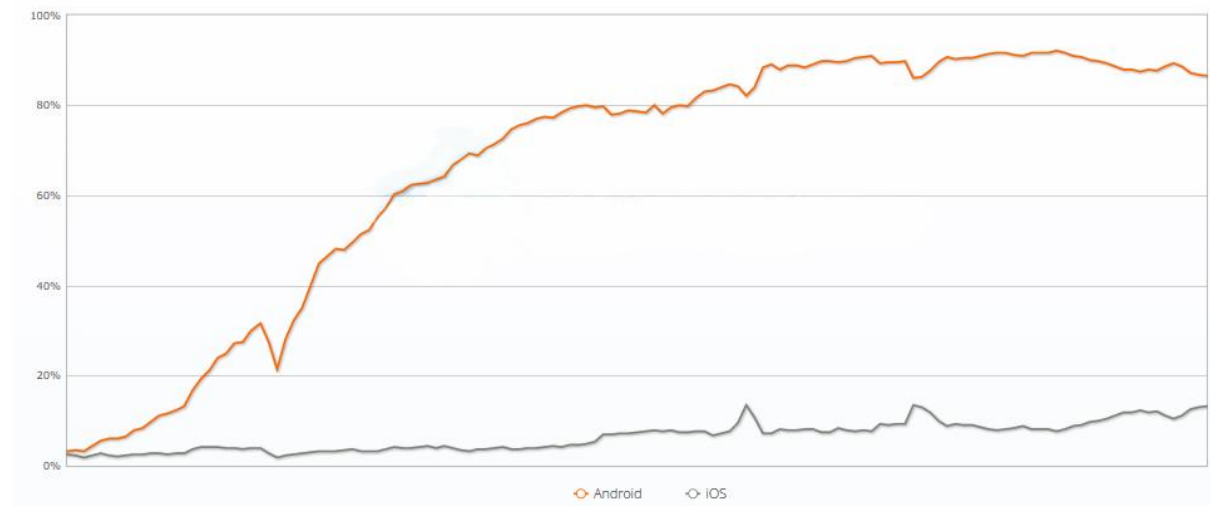


Figure 1 : Android and IOS OS Market in Sri Lanka

Due to the open-source nature of the platform, which enables developers to build specialized applications suited to requirements, Android-based apps have grown in popularity among developers. Developers can build apps that are easier for people who are blind to use thanks to this flexibility. To enhance the user experience, the application might incorporate disability features such as text-to-speech capabilities.

Android-based applications offer a high degree of customization, allowing developers to create applications that meet the specific needs of visually impaired individuals in Sri Lanka. For example, developers can create applications that support the Sinhala language and integrate text-to-speech features that improve the user experience.

The large developer community for Android offers tools so that programmers can make apps that are usable by people with vision impairments. Developers can use these materials to build mobile book devices that support the Sinhala language and accessibility features. These resources also include open-source frameworks and tools.

2. Research Gap

According to the literature review, there are several mobile applications for Text-to-Speech (TTS). Researchers have used a variety of methods, such as image processing and machine learning, to recognize writing and turn it into voice. To satisfy the requirements of Sri Lankans, there are currently no apps accessible that are specially designed with Sinhala Text-to-Speech (TTS) capabilities. For people and groups who depend on these services daily, this can be a significant restriction. But the available Text-to-Speech (TTS) mobile apps for the Sinhala language are limited and only a few are included in the table below.

Table 1 : Comparison Between Existing Systems

Application Reference	Research A	Research B	Research C	Research D	Our Sinhala Book Reader
OCR send to captured text Text-to-Speech (TTS) Synthesizer in real time	✗	✓	✗	✗	✓
Text Preprocessing to extract Meaningful information.	✓	✗	✗	✓	✓
Use cloud / diphone database to Store the Data	✓	✓	✗	✗	✓
Adjust the speed of the voice output.	✗	✗	✗	✗	✓
Support For Multiple Languages	✓	✗	✓	✓	✗
Android Mobile Application	✗	✗	✗	✗	✓

Natural language processing (NLP) is a component of a text-to-speech (TTS) generator known as the front end. The task of creating a symbolic depiction of the text to be considered depends on this NLP function. This depiction combines the words' phonetic transcription along with the intended tone and cadence, or prosody. The NLP module is followed by a digital signal processing module (DSP, also termed as backend), which transforms the symbolic information it receives into speech [3].

3. Research Problem

The World Health Organization (WHO) estimates 285 million people are visually impaired, including 39 million blind, globally [5]. For those who are visually impaired individuals, reading books can be very difficult because they commonly cannot obtain printed materials in a way that is convenient for them. Despite improvements in assistive technology like text-to-speech software and Braille screens, many books are still not widely accessible in a manner that the blind can read. For blind people who want to increase their knowledge, participate in literary encounters, and enhance their educational and job possibilities, this restricted accessibility can present significant hurdles.

Another problem is the price of specialist equipment and software, which for many blind individuals can be unaffordable. Even if these tools are accessible, perusing a conventional paper book might not be a similar experience with them. For instance, braille displays can be clumsy and sluggish, and text-to-speech applications can have trouble with complicated English and formatting.

The lack of audiobooks and Braille resources is another problem. Even though there are more audiobooks available, there are still few of them compared to the enormous number of written novels. Due to the lengthy and expensive procedure involved in converting written books into Braille, finding Braille books is even more difficult. This implies that the most recent bestsellers or widely used instructional resources may not be available to blind people.

In conclusion, a shortage of access to written materials is the primary obstacle that blind people encounter when perusing books. Even with advancements in assistive technology, there are still many obstacles to be overcome, including the price of specialized hardware and software, the scarcity of audiobooks and Braille materials, and the challenge of providing a reading experience that is comparable to that of conventional printed books. To overcome these obstacles, there must be a concerted effort to increase the accessibility of books for blind readers and to guarantee that they have the same access to reading and educational possibilities as sighted people.

4. Objectives

4.1. Main Objective

A Sinhala book reader for the visually impaired is a software program designed to make reading accessible for individuals with visual impairments. It combines various technologies to provide a seamless reading experience.

Here, the Text-to-Speech (TTS) system and an Optical Character Recognition (OCR) system to recognize the frequently used Sinhala characters is the two major components of the software that we plan to create [6]. The device utilizes optical character recognition (OCR) technology to convert the text from a physical book into a digital format. Then, a text-to-speech synthesizer reads the text out loud in the Sinhala language, making it easier for visually impaired users to follow along.

In addition to the text-to-speech synthesizer, the device also includes audible guidance to help navigate the app and identify the distance to the book being read. This makes it easier for visually impaired users to find their place in the book and keep track of their progress.

4.2. Sub Objective

4.2.1. Text-to-Speech (TTS)

Speech synthesis is the term used to describe the mechanical production of human-like speaking. A speech synthesizer is a computerized device created to produce artificial sounds that resemble real speech. Text to Speech converts language text into speech whereas other systems convert representations of symbolic linguistics into speech [2].

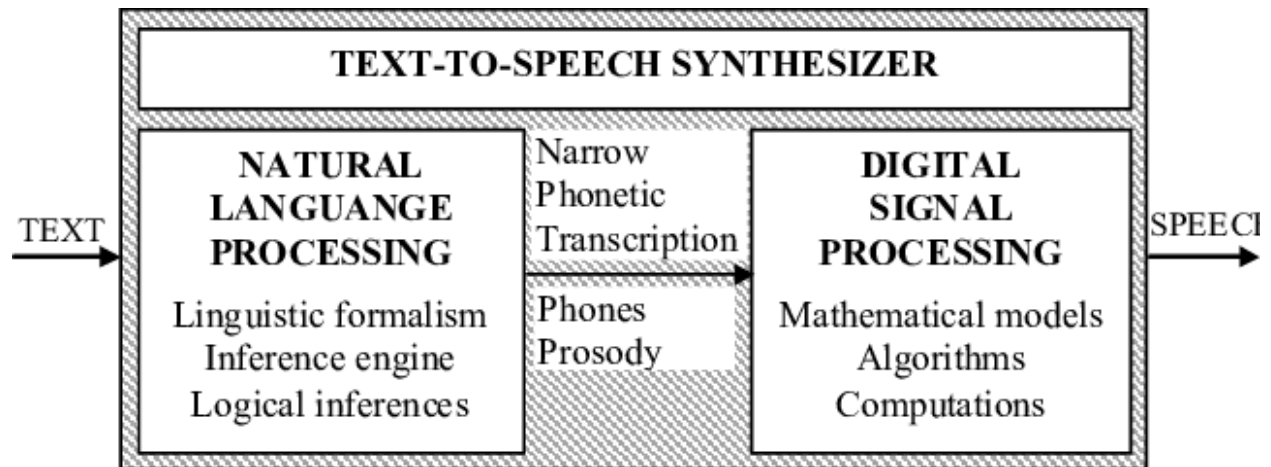


Figure 2 : General Functional Diagram for Text-to-Speech (TTS) System [7]

Concatenation, which entails combining pre-recorded speech fragments kept in a database, is a popular technique for producing artificial speech around speech synthesis. Another strategy includes turning written text into speech with the Text-to-Speech (TTS) method, and some systems use various strategies to change symbolic language representations into voice. Speech synthesis uses two main technologies formant and concatenation to create synthetic speech that closely resembles natural speech.

5. Methodology

5.1. Overall System Diagram

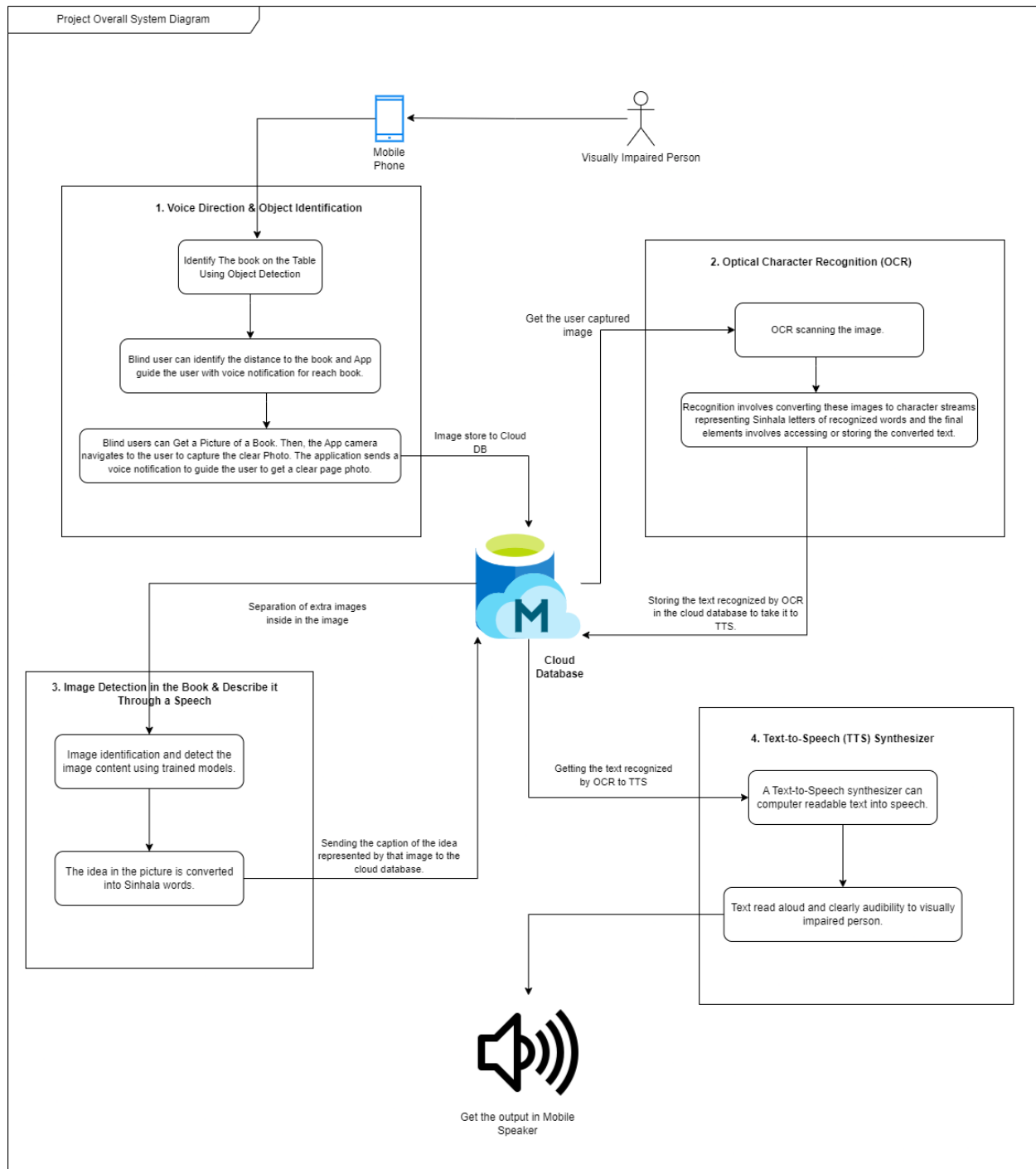


Figure 3 : Project Overall System Diagram

Audible guidance helps the user navigate the app through the functions the app and get a clear idea and guidance whenever the user faces difficulty performing a task. When a user wants to find a book on the table, he opens the camera through the app and points it toward the table or desk and the app identified the user's hand and navigates to the book. Using real-time Image processing technology identifying the dangerous object near the blind user and the distance to the object will notify verbally to the user and identify the probability of occurring an accident the user. Users will navigate in a pristine environment evading dangers and harmful things.

Sinhala Character Identification and word formation through the engine and translate to the speech then send to the TTS. When the app is not running, users can read the time using a background process. The program should also be able to start the camera when the user launches it using voice commands. The program should be able to quickly scan the document when the user launches it. On the paper in front of the camera, the app needs to be able to automatically focus. Until the document is within the capture frame, the system should alert the user audibly. When the user wants to capture an image on the book the app alerts and navigates the user to capture the image onto the frame of the phone. The device's storage should be used to store the image that was captured. Before submitting information to the OCR system, the system must detect and correct skew.

A computer vision system that uses cutting-edge methods to recognize and describe objects and scenes in real-time is an image detection software for blind students. The program takes pictures of the user's surroundings using the camera on a smartphone or tablet. The objects and their characteristics are then identified from these photos using image processing techniques like edge detection, color analysis, and feature extraction. After locating the things in the image with the use of object detection algorithms, machine learning models are used to identify and categorize the objects. These models use methods like convolutional neural networks (CNNs) to recognize objects in real-time and are trained on massive datasets of annotated photos.

Here, the main purpose of using Text to speech (TTS) technology is to give a blind person the ability to access the written text of a Sinhala book. This allows them to easily listen to the valuable content of Sinhala books. TTS technology allows the written text in a Sinhala book to be read out loud in a natural-sounding voice, which makes it easier for visually impaired people to understand the content. The technology uses computer algorithms to analyze the Sinhala text and generate an appropriate pronunciation, intonation, and rhythm for each word and sentence.

5.2. Individual System Diagram

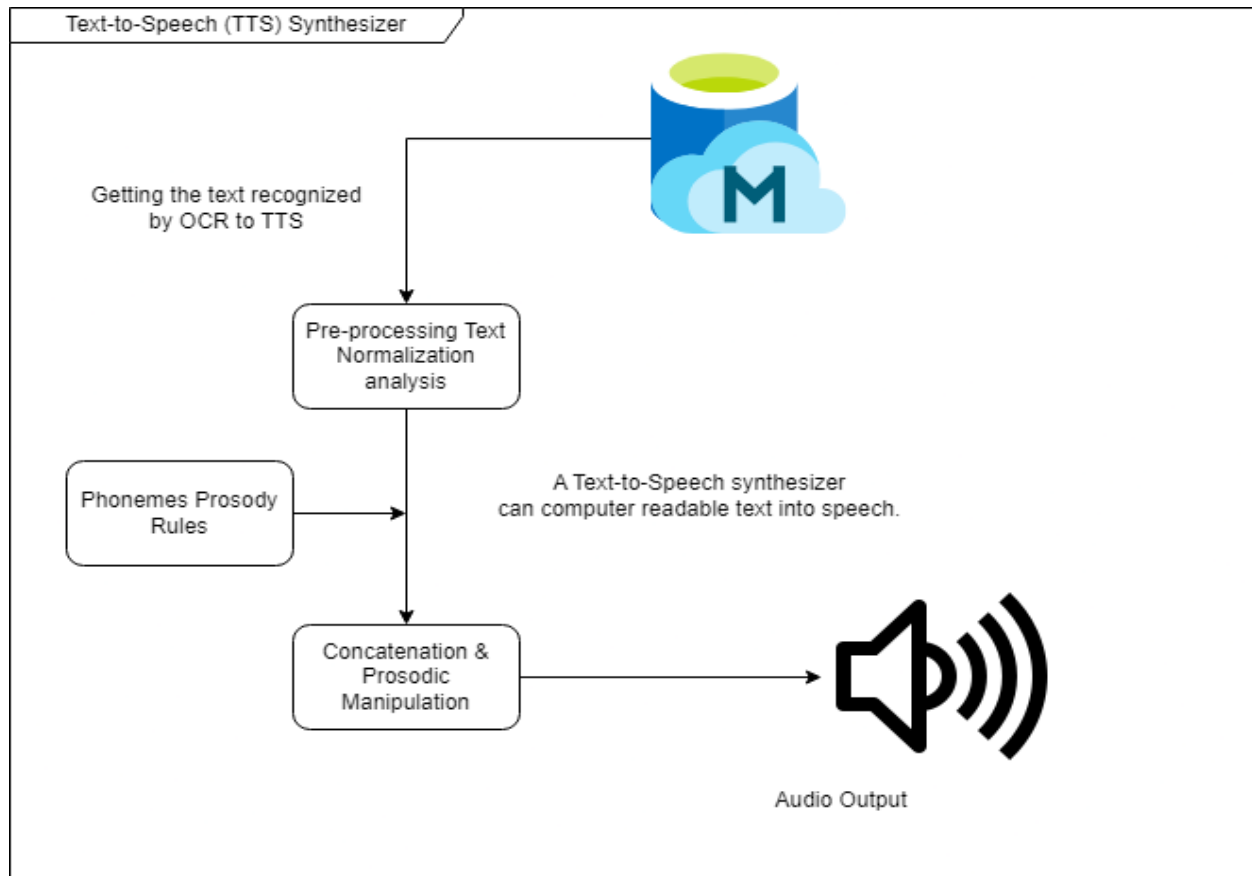


Figure 4 : Text-to-Speech (TTS) Synthesizer

A Text-To-Speech synthesizer is a computer-based system capable of converting computer-readable text into speech. The conversion of text to speech involves many important processes. These processes can be divided mainly into three stages; text analysis, linguistic analysis, and wave-form generation [4].

Here are the steps,

- An audio notification should be sent to a visually impaired person after the OCR process is complete.
- Here, Texts that have been extracted and identified in Sinhala characters should be transmitted to the TTS system.
- Finally, TTS should be used to read aloud and clearly audibly to visually impaired people the Sinhala text recognized in the camera picture.
- Overall, the main task of using TTS for visually impaired people in a Sinhala book reader is to provide them with a way to access written text and enjoy the content of books, regardless of their visual impairment.

6. Project Requirements

6.1. Functional Requirements

- **Text-to-Speech (TTS) Functionality:** The book reader should be able to use TTS to read the content audibly in Sinhala.
- **Offline Reading:** Reading books while not connected to the internet is possible with the reader's offline reading feature, which enables users to acquire and keep books for offline reading.
- **Voice Controlling:** To enable users to use vocal commands to navigate the application and the book, the reader ought to have a voice control function.
- **Easy Navigation:** Users should have numerous choices for adjusting the speed and pitch of the TTS speech, as well as big buttons and signs that are simple to read.

6.2. Non-Functional Requirements

- **Accessibility:** The program needs to be created with people who are visually impaired individuals in mind.
- **Usability:** The program must be simple to use, have simple controls, and provide users with straightforward directions.
- **TTS voice quality:** The application's Text-to-Speech (TTS) engine must generate high-quality, believable speech.
- **Performance:** Even on mid-range devices, the program must be prompt and quick.

7. References

- [1] WHO, *World report on vision*, vol. 214, no. 14. 2019. [Online]. Available: <https://www.who.int/publications-detail/world-report-on-vision>
- [2] A. Mishangi, "Android based sinhala document reader for visually impaired people," 2021.
- [3] R. Carlson and B. Granström, "Speech Synthesis," *Handb. Phonetic Sci. Second Ed.*, pp. 781–803, 2010, doi: 10.1002/9781444317251.ch21.
- [4] K. Wasala, A., Weerasinghe, R., Gamage, "A Sinhala Text-to-Speech System," 2007.
- [5] G. V. S. Murthy *et al.*, "The Sri Lanka National Blindness, Visual Impairment and Disability Survey: rationale, objectives and detailed methodology," *Ceylon Med. J.*, vol. 63, no. 5, p. 3, 2018, doi: 10.4038/cmj.v63i5.8735.
- [6] R. Weerasinghe, A. Wasala, D. Herath, and V. Welgama, "NLP applications of Sinhala: TTS & OCR," *IJCNLP 2008 - 3rd Int. Jt. Conf. Nat. Lang. Process. Proc. Conf.*, vol. 2, pp. 963–966, 2008.
- [7] N. A. Azis, T. V. Tjahja, and A. S. Nugroho, "Evaluation of Text-to-Speech Synthesizer Unpredictable Sentences Test : IndoTTS , eSpeak , and," no. April 2014, 2011.