BLIND READER

Therese Johnson[1], Justina Babu[2], Rismy C. J[3], Dr.Rajeswari M[4]
[1,2,3] U.G.Scholar, [4] Associate Professor,
[1,2,3,4]Department of Computer & Engineering,
[1,2,3,4]Sahrdaya College of Engineering and Technology,Kodakara,Kerala,India.

Abstract

Human communication today is mainly via speech and text. To access information in a text, a person needs to have vision. However, those who are deprived of vision can gather information using their hearing capability. The proposed method is a camera based assistive text reader to help blind person in reading. The proposed project involves Text Extraction from image and converting the Text to Speech using text to speech convertor, a process which makes blind people to collect information given in the text. Both the printed text and handwritten text can be made accessible for the blind people using this project. This project also aims on translating the text into some specified languages opted by the user. This project is carried out in open CV using the language C++. Translation will be done using python.

Keywords: Optical Character Recognition (OCR), Text To Speech (TTS), Tessarac, Artificial Neural Network (ANN), Machine Learning, Support Vector Machine(SVM).

INTRODUCTION

In our society there are many individuals with sensory disabilities like blindness, learning disabilities, language or cultural differences, etc and they all may require different ways of approaching content. Our project is to help such Individuals to gather knowledge .It also enables visually or physically challenged, senior citizens and even the illiterate people to get benefit of reading confidential documents without depending on others. Our project is developed with a vision of empowering and enabling these Individuals. Here the system consists of a camera and a laptop which acts as an interface between the system and the user. With some conditions, the image is captured and then it will recognize character, which are then read out by the system through a speaker and if the user needs translation, it can be translated to opted language. Here the image of both handwritten and printed text can be processed. It speaks out the content that was converted in to text format in the system from processing the image of the paper. So the system provides audio output to user. This helps both illiterate and blind people to understand the content of their documents without the help of others. Converting text to speech, helps the visually challenged to read and understand various documents. This is also helpful if a person is very busy and he wants to do multitasking then he can listen to the audio outputted and at the same time he can do some other works.

There are many advantages such as it can Prevents eye from strain, and user can sit and listen comfortably, saves time especially while driving, exercising, easy to use, improves reading, also a student who have dyslexia is able to write a sentence and then listen to their writing. This enables them to decide if what they have composed is correct. This is very helpful for building better speech and literacy skills. There are people with learning disabilities – Some people have difficulty reading large amounts of text due to dyslexia and other learning disabilities. This project offers them an easier option for experiencing content. Also helps people who have literacy difficulties – Some people have basic literary levels. They often get frustrated trying to read large documents because so much of it is in text form. By offering them an option to hear the text instead of reading it, they can get valuable information in a way that is

more comfortable for them, People who speak the language but do not read it – Having a speech option for the foreign born will open up your audience to this under-served population.

Many people who have come to a new country learn to speak and understand the native language effectively, but may still have difficulty in reading and writing the same language. Though they may be able to read content with a basic understanding, our project allows them to take in the information in a way they are more comfortable with, making the content easier to comprehend and retain. People who have multitask – A busy life often means that people do not have time to do all the reading. Having a chance to listen to the content instead of reading it allows them to do something else at the same time. People with visual impairment – Text to speech can be a very useful tool for the mild or moderately visually impaired. Even for people with the visual capability to read, the process can often cause too much strain to be of any use or enjoyment.

With text to speech, people with visual impairment can take in all manner of content in comfort instead of strain. People with different learning styles – Some people are auditory learners, some are visual learners, and some are kinesthetic learners – most learn best through a combination of the three. Universal Design for Learning is a plan for teaching which, through the use of technology and adaptable lesson plans, aims to help the maximum number of learners comprehend and retain information by appealing to all learning styles .Here our project attempt to convey what people with normal eyesight see on a display to their users via non-visual means, like text-to-speech , and it provides hands free access to the printed and handwritten documents.

METHODOLOGY

In this system , first the image of the required text document is captured using a high resolution camera. Both the printed and handwritten text image can be captured. We build databases for the handwritten documents and we then train the system using machine learning. The next important step is that we extract text from the captured image using OCR. Extracted text is then converted to speech using the TTS system. This is the main aim of our project. Translation is also made possible through this project ie, we can translate the text to some specified language requested by the user and then it is again converted to speech.
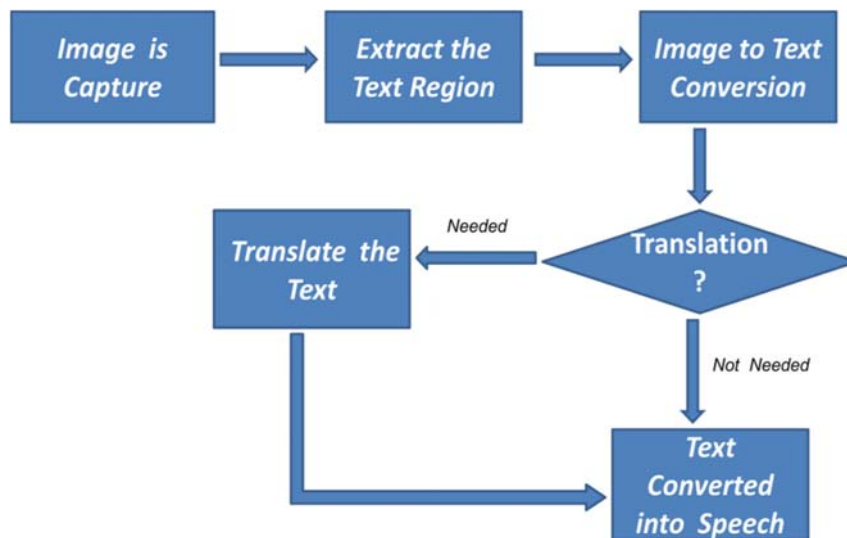
Block Diagram of Proposed System

Figure .1

Many researches are going on in the text to speech conversion field. Speech synthesis is the main feature we are using in this project. Speech synthesis is the artificial production of human speech. A computer system used for this purpose is called a speech computer or speech synthesizer, and can be implemented in software or hardware products. A text-to-speech (TTS)

system converts normal language text into speech; other systems render symbolic linguistic representations like phonetic transcriptions 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.

This project is actually intending on converting both the printed and handwritten text to speech. Translation is also made possible by this project. Printed text documents can be converted more easily than converting handwritten texts. For converting handwritten text we must make more number of databases which increases the efficiency. The main modules in our project are database building, image capturing, image to text conversion, text to speech conversion, translation.

DATABASE BUILDING

A database is an organized collection of data. A database is not generally portable across different DBMSs, but different DBMSs can interoperate by using standards such as SQL and ODBC or JDBC to allow a single application to work with more than one DBMS. Computer In this module, we collect databases as much as possible for getting maximum efficiency in case of handwritten text conversion. We know that handwriting of each and every person is different. So we must create database of maximum letters, numbers etc and we must train the system using the technique machine learning. Machine learning is a field of computer science that gives computers the ability to learn without being explicitly programmed. Machine learning is closely related to (and often overlaps with) computational statistics, which also focuses on prediction-making through the use of computers. It has strong ties to mathematical optimization, which delivers methods, theory and application domains to the field.

IMAGE CAPTURING

Image Capture is an application program from Apple that enables users to upload pictures from digital cameras or scanners which are either connected directly to the computer or the network. It provides no organizational tools like iPhoto but is useful for collating pictures from a variety of sources with no need for drivers. We are making use of this same technique in our project. We are using a high resolution camera to capture the required document image.

IMAGE TO TEXT CONVERSION

Here we are converting the captured image to text by using OCR(optical character Recognition). Optical character recognition (also optical character reader, OCR) is the mechanical or electronic conversion of images of typed, handwritten or printed text into machineencoded text, whether from a scanned document, a photo of a document, a scene-photo (for example the text on signs and billboards in a landscape photo) or from subtitle text superimposed on an image (for example from a television broadcast). It is widely used as a form of information entry from printed paper data records, whether passport documents, invoices, bank statements, computerised receipts, business cards, mail, printouts of static-data, or any suitable documentation. It is a common method of digitising printed texts so that they can be electronically edited, searched, stored more compactly, displayed on-line, and used in machine processes such as cognitive computing, machine translation, (extracted) text-to-speech, key data and text mining. OCR is a field of research in pattern recognition, artificial intelligence and computer vision.

TEXT TO SPEECH CONVERSION

Text to speech (TTS) is the use of software to create an audio output in the form of a spoken voice. The program that is used by programs to change text on the page to an audio output of the spoken voice is normally a text to speech engine. TTS engines are needed for an audio output of machine translation results. A text-to-speech system (or "engine") is composed of two parts: a frontend 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, pre-processing, 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-tophoneme 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.

TRANSLATION

Translation is the communication of the meaning of a source-language text by means of an equivalent target-language text. Computer-assisted translation is a broad and imprecise term covering a range of tools, from the fairly simple to the complicated. These can include:

- Translation memory tools (TM tools), consisting of a database of text segments in a source language and their translations in one or more target languages.
- Spell checkers, either built into word processing software, or add-on programs.
- Grammar checkers, again either built into word processing software, or add-on programs
- Terminology managers, which allow translators to manage their own terminology bank in an electronic form.
- Full-text search tools (or indexers), which allow the user to query already translated texts or reference documents of various kinds. Some such indexers are ISYS Search Software, dtSearch Desktop and Naturel.
- Bitext aligners: tools that align a source text and its translation which can then be analyzed using a full-text search tool or a concordance.

FINDINGS

Now a days many text to speech conversion devices are already available. These devices make use of a scanner which scans the entire document and subsequently the in-built optical recognition system extracts the text from the page. There are different technologies like Top-Braille, Trinetra and speech synthesis, screen reader, finger reader, Braille notetaker and printer,e-bookreader or audio book reader etc. In this project we are capturing the image using a camera and using image extraction features we are extracting the text. Basically handwritten text conversion is a difficult task since the handwritings of every person is different. In this project, we are converting the handwritten text to speech as well as the printed text. Translation of some form of text into desired language is another attractive feature in this project. All these features make this project unique from all the existing systems.

CONCLUSIONS

The 'blind reader' is not just a project that empowers the blind and illiterate people to become independent, but is also a resource saver. This project is an effort to implement an innovative robust approach for character extraction and text to voice conversion of different images using optical character recognition. A user friendly, cost effective, reliable to all and applicable in the real time system is achieved. By specifying printed or handwritten text the system can processes it. Using this methodology, we can read text from a document, whether it is printed or handwritten and can generate synthesized speech through any portable system i.e. computer's speakers. This saves user's time by allowing him to listen background materials while performing other tasks. Other application of this system includes such as making information browsing for people who do not have the ability to read or write. This approach can be used in part as well. If requirement is only for text conversion then it is possible or else text to speech conversion is also done easily. People with vision impairment or visual dyslexia or complete blindness can use this approach for reading the documents, books and also while travelling. Tourists having language barrier can also use this approach for understanding different languages. Translation to the user opted language is also make possible. People travelling in cars and buses can save their time using this feature. Experiments have been performed to test the text and speech generation system and good results have been achieved. Still the work is in progress for symbol extraction.

REFERENCES

1.Bazzi, I.; Schwartz, R.; Makhoul, J. (Jun 1999), An Omni font openvocabulary OCR system for English and Arabic, Pattern Analysis and Machine Intelligence, IEEE Transactions on vol.21, no.6, pp.495-504.

2.Devika Sharma ,Ranju Kanwar ,” Text To Speech Conversion With Language Translator Under Android Environment ” Publish In International Journal Of Emerging Research In Management &Technology Issn: 2278-9359(Volume-4,Issue-6).

3. Kalyani Mangale, Hemangi Mhaske, Priyanka Wankhade, Vivek Niwane, “Printed Text *To Audio Converter Using OCR”, International Journal Of Computer Applications (0975*–8887) (NCETACt2015).

4. Pardeep Kaur, Pooja Choudhary, Varsha Sahni, “English scanned document character *recognition and matched and missed matched analysis using NN and MDA”, *Inteernational Journal of computer Applications,2015.

5. Smith, R. (2007), An Overview of the Tesseract OCR Engine. In proceedings of Document analysis and Recognition. ICDAR 2007. IEEE Ninth International Conference.

6. Suraj Mallik, Rajesh Mehra,”Speech To Text Conversion For Visually Impaired Person Using Â Law Companding ” Publish In Iosr Journal Of Electronics And Communication Engineering (IosrJece) E-Issn: 2278- 2834,P- Issn: 2278-8735.Volume 10, Issue 6, Ver. Ii (Nov - Dec .2015).

7. V. Ajantha Devi, Dr. S Santhosh Baboo (Jul-Aug 2014), Optical Character Recognition on Tamil Text Image Using Raspberry Pi International Journal of Computer Science Trends and Technology (IJCST) – Vol. 2 Issue 4.