International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 06 Issue: 05 | May - 2022 Impact Factor: 7.185 ISSN: 2582-3930

# SPEECH SYNTHESIS BASED ON THE RECOGNITION OF OPTICAL CHARACTER USING LABVIEW

**Kamali.K[1], Prof Dr.Azha.Periasamy[2],**

[1]*Department of Electronics and Instrumentation & Bharathiar University, Coimbatore 46,Tamil Nadu,India*
[2]*Department of Electronics and Instrumentation & Bharathiar University, Coimbatore 46, Tamil Nadu,India*

---

**Abstract –** The speech synthesis of voice signal utilizing labVIEW with text and signal input in the form of picture or pdf is shown in this study. Optical character recognition is used to synthesize the signal. It integrates OCR and speech synthesis into one system. Speech synthesis is quite beneficial in the classroom.

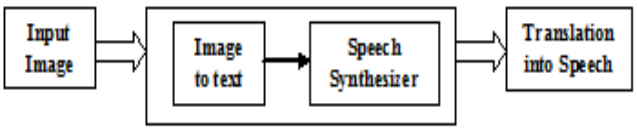
***Key Words*:** OCR(optical character recognition), speech synthesis, LabVIEW

## INTRODUCTION

The acronym OCR stands for optical character recognition. It's the process of converting scanned images of handwritten or printed text into machine-readable text by mechanical and technological means. It transforms text into a vocal signal. Synthesis is the process of creating a speech signal from a text or image file. Text to speech is another name for speech synthesiser. This is the most natural and effective way for human beings to communicate.

## OCR

Optical Character Recognition (OCR) is an acronym for optical character recognition. The above technology uses an optical system to recognize characters automatically. Our optics, in the case of humans, are optical mechanisms. The image viewed by the eyes serves as input to the brain. Many factors influence each individual's capacity to comprehend these inputs. OCR is a technology that simulates human reading ability. Although OCR cannot compete with the reading ability of humans.




OCR

Both handwritten and printed word can be recognized using OCR. However, the quality of the given document has a direct impact on OCR performance. OCR is a type of image processing software. It almost completely consists of text, with very little non-text clutter obtained from a photo taken with a mobile camera. This program that combines Google's open-source OCR engine, Tesseract, with a text recognition OCR engine for the Android system. OCR is a technology that allows the user to enter multiple forms of documents into editable and usable data, including such scanned paper documents, PDF files, or digital camera photographs. Digital camera images are not the same as scanned documents or images.

They typically have shortcomings like edge distortion and weak illumination, making it difficult for most OCR software to scan the text correctly. Receive more information was chosen because to its widespread acceptance, extensibility and flexibility, active development community, and the fact that it "simply works" out of the box. Our application must go through three crucial steps in order to conduct character recognition. The first is segmentation, which involves identifying distinct glyphs (basic units encoding one or more characters, usually continuous) from a binary input image. The second phase is feature extraction, which involves computing from each data point.

## TEXT TO SPEECH

A text - to - speech (TTS) synthesizer is a system that can automatically read text aloud that has been extracted from Optical Character Recognition (OCR). Both hardware and software can be used to implement a speech synthesizer. The artificial production of human speech is known as speech synthesis. A speech synthesizer is a computer system that is used for this purpose. Text-to-speech (TTS) systems convert written text into speech. To create a completely "synthetic" voice output, a synthesizer can incorporate a model of the vocal tract and other human voice characteristics.

## SYNTHESIS OF SPEECH

The artificial production of human speech is known as speech synthesis. Synthesizing is a highly efficient method of generating speech waveforms using machines based on phonetical transcription of the message. Recent advances in speech synthesis have resulted in synthesisers with extremely high intelligibility, but sound quality and naturalness remain major issues.

## EXPERIMENTAL SETUP

The mechanical or electronic conversion of images of handwritten or printed text into machine-editable text is known as optical character recognition (OCR). The process steps for the OCR-based system are as follows:

Image Capture, Image Pre-processing (Binarization), Image Segmentation, Matching and Recognition.

## IMAGE CAPTURE

The image was captured using a digital HP scanner. To achieve a uniform black background, the scanner's flap had been left open during the acquisition process. The image was captured with the help of a Lab VIEW programmer. The Image was configured with the help of the Image to create a LabVIEW sub-function. The image configuration entails selecting the image type and border size based on the requirements. This workmen8-bit image was used with a border size of 3.

## PRE-PROCESSING OF IMAGES (BINARIZATION):

Binarization is the process of converting a grayscale image (0 to 255 pixel values) to a binary image (0 to 1 pixel values) by applying a threshold value. The pixels that were lighter than the threshold were returned to white, while the rest were converted to black pixels. In this work, global thresholding with a threshold value of 175 was used to binarize the image, which means that pixels with grayscale values between 175 and 255 were converted to 1, while pixels with grayscale values less than 175 were converted to 0.

## IMAGE SEGMENTATION

Line segmentation, word segmentation, and finally character segmentation comprise the segmentation process.

## LINE SEGMENTATION

The first step in the segmentation process is line segmentation. It takes the image's array as input and scans it horizontally to find the first ON pixel, remembering the coordinates as y1. Since the characters would have begun, the system has continued to scan the image horizontally and discovered many ON pixels. When the first pixel is detected, the system remembers the coordinates as y2 and checks the pixel's surroundings to determine the required number of OFF pixels.

## WORD SEGMENTATION

The line segmented images were vertically scanned to find the first ON pixel during the word segmentation process. When this occurs, the system remembers this point's coordinate as x1. This is the word's starting coordinate. The scanning process is repeated until fifteen (assumed word distance) consecutive OFF pixels are obtained. The first OFF pixel is recorded as x2 by the system. The phrase is from x1 to x2. All of the words were segmented in this manner, and the segmented words were used in the next step for character recognition.

## CHARACTER SEGMENNTATION

By scanning the word segmented image vertically, character segmentation was performed. This procedure differs from word segmentation in the following two ways:

I The number of horizontal OFF pixels between different characters is less than the number of OFF pixels between words.

ii) The total number of characters and their order in the word have been determined in order to correctly reproduce the word during speech synthesis.

## MATCHING AND RECOGNITION

The correlation VI determines the relationship between segmented characters and each character's stored templates. The highest correlation value identifies a specific character. To recognize the character, each segmented character was compared to predefined data stored in the system. Because the same font size was used for recognition, each character has a unique match.

## TEXT TO SPEECH SYNTHESIS

In-text-to-speech module text recognized by OCR system will be the inputs of speech synthesis system to be converted into speech in.wav file format and creates a wave file named output wave, which can be listened to using wave file player. Text to speech synthesis consists of two steps:

i)Text to speech synthesis

ii) Play a speech in. wave format.

## TEXT TO SPEECH CONVERSION

Using Automation open, invoke node, and property node, input text is converted to speech (in LabVIEW). A LabVIEW program me for text-to-speech conversion is shown, as well as a flow chart for text-to-speech conversion. Characterizes the flow chart for playing the converted speech signal, while illustrates its LabVIEW programmers.

## OUTCOME AND DISCUSSION

The OCR-based speech synthesis system is divided into two parts. There are two of them:

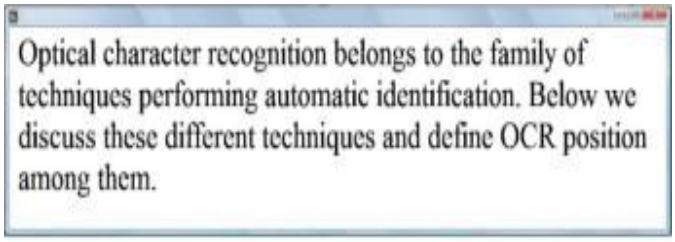
i)optical character recognition

ii)speech to synthesis

## OPTICAL CHARACTER IDENTIFICATION

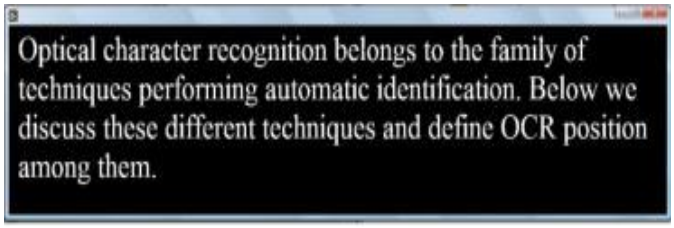
**s**tep 1:

The scanner scans the printed text, and the system uses IMAQ to read the image. Read the file and display the image using the LabVIEW IMAQ Wind Draw function.




step 2

The image was binarized with a threshold of 175 in this step. and the resultant image has been displayed




Step 3

In this step line segmentation of threshold image has been done.


Step 4

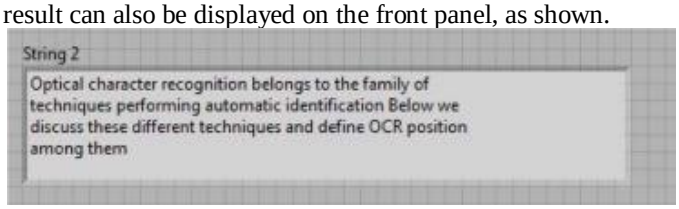
Words have been segmented from the line in this step.


Step 5

In this step, character segmentation was performed, and all of the characters in the word image window were segmented. The segmentation of the first three characters of the word "Optical" has been demonstrated. Following segmentation, each character was correlated with stored character templates, and the system recognized printed text.

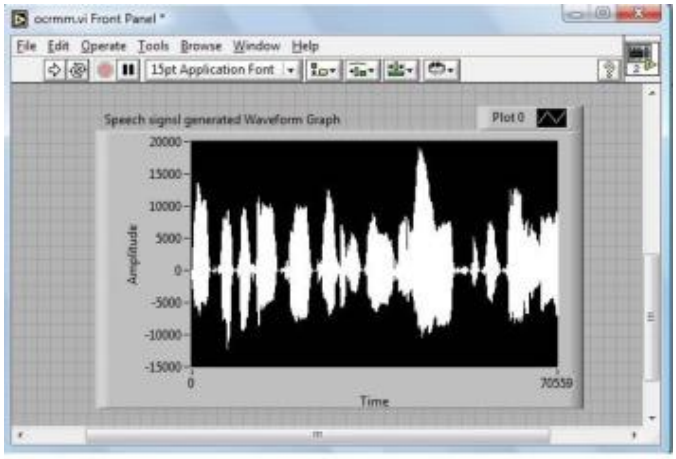
Step 6

Finally, the OCR system output is in text format, which is saved in a computer system. The recognized text result can also be displayed on the front panel, as shown.




## SPEECH SYNTHESIS:

A wave file output.wav is created containing text converted into speech which can listen using wave file player. The waveform will vary according to the different text from OCR output in the text box and can be listened on the speaker. The wave form for above recognize text has been shown




## CONCLUSION

A speech synthesis system based on OCR (can be used as a good mode of communication between people). OCR and speech synthesis are components of the developed system. IMAQ Vision for LabVIEW was used to scan printed or written character documents and acquire images for OCR. The various characters were identified using segmentation and correlation methods developed in LabVIEW. The recognized text has been converted into speech in the second section. The developed OCR-based speech synthesis system is simple to use, inexpensive, and provides results in real time. This library allows for the selection of the voice and sound device to be used, the OCR perceived content to be read, and the modification of the rate and volume of the selected voice.

## FUTURE SCOPE

OCR-based Speech Recognition Framework Using LabVIEW is a skilled programmers that produces excellent results for specific textual styles (equivalent to or greater than 48 text dimension), but there are still opportunities to improve it. The framework can be improved by using the Omni font. OCR-based speech recognition framework created in LabVIEW. There is a decent future level to make it for utilizing various techniques all the more quickly and productively. Our next works with OCR Mobile Application will include changing the outcomes by utilizing table limits discovery procedures and utilizing content post-preparation Techniques to distinguish the clamour and to redress poorly perceived words. The OCR application will also display the marks and alternate images as they appear in the report.

## REFERENCE

**1. Wei, T. C., Sheikh, U. U., & Ab Rahman, A. A. H**. (2018, March). Improved optical character recognition with deep neural network. In *2018 IEEE 14th International Colloquium on Signal Processing & Its Applications (CSPA)* (pp. 245-249). IEEE.

**2. Mullani, J. J., Sankar, M., Khade, P. S., Sonalkar, S. H., & Patil, N. L.** (2018, February). OCR based speech synthesis system using labview: text to speech conversion system using OCR. In *2018 second international conference on computing methodologies and communication (ICCMC)* (pp. 7-14). IEEE.

**3. Singla, S. K., & Yadav, R. K.** (2014). Optical character recognition based speech synthesis system using labview. *Journal of applied research and technology*, *12*(5), 919-926.

**4. Chinnaiah, M. C**. Low cost Speech Synthesis System Using Virtual Instrumentation and Arduino.

**5. Bansal, P**. (2013). An Interactive Speech Synthesis System Based on OCR Using labview.

**6. Singla, S. K., & Yadav, R. K**. (2014). Optical character recognition based speech synthesis system using labview. *Journal of applied research and technology*, *12*(5), 919-926.

**7. Patil, A. V., & Mehta, P**. (2016). An Accurate and Revised Version of Optical Character Recognition Based Speech Synthesis Using Labview.

**8. Sharma, R., Singla, S. K. G., & Singh, N. J. G.** (2007). *Speech Synthesis* (Doctoral dissertation).

**9. Balaji, B. A., & Mohan, R. R.** Optical character based tts synthesizer using myrio