Volume-4, Issue-8, August-2015 • ISSN No 2277 - 8160

Research Paper Engineering

Text to Speech Synthesis for Blind Person

Yamini D.Patil Student Of M.E. (E&TC), E& TC Dept., D.N.Patel Collage Of Engg.,Shahada, Maharashtra

Nandakishor C.Patil Prof. Of E&TC Dept, D.N.Patel Collage Of Engg.,Shahada, Maharashtra

ABSTRACT

The present growth of digitization of books and manuscripts demands an immediate solution toaccess them electronically. In the last three decades significant advancement is made in therecognition of documents written in Latin-based scripts. There are many excellent attempts inbuilding robust document analysis systems in industry, academia and research labs. While in textto speech, there are many systems which convert normal language text in to speech. This thesisaims to study on image recognition technology (Optical Character Recognition) with speechsynthesis technology and to develop a cost effective user friendly image to speech conversionsystem using MATLAB. In this work we tried to make a system by which we can get the textthrough image and then speech through that text using MATLAB. The primary motivations are

to provide users with a friendly vocal interface with the computer and to allow people withcertain handicaps (such as blindness, dumbness, poor vision, visual dyslexia) to use the computeror to read any type of documents.

KEYWORDS : OCR,MATLAB, Speech synthesis, speech conversion, Character Extraction algorithm.

INTRODUCTION

Navigation aid helps the visually-impaired, or those working in low visibility areas, navigate indoors by assisting in the identification and avoidance of obstacles. The typical support instrument for the visually impaired is the "white cane", which is used for the detection of obstacles by providing a limited degree of feedback to the user. The portable indoor navigation aid alerts the user to the layout of an entire indoor area upon entrance so as to assist in collision avoidance. Now a days in markets number if navigation aid are available for Blind persons.

Our main intension is to provide a system through which blind person can get the knowledge of words they cant see. Here we use Optical character recognition & speech synthesis technique.

OCR is the acronym for Optical Character Recognition. This technology allows a machineto automatically recognize characters through an optical mechanism[2]. Human beingsrecognize many objects in this manner our eyes are the "optical mechanism." But while thebrain "sees" the input, the ability to comprehend these signals varies in each personaccording to many factors. By reviewing these variables, we can understand the challenges.

First, if we read a page in a language other than our own, we may recognize the variouscharacters, but be unable to recognize words. However, on the same page, we are usuallyable to interpret numerical statements - the symbols for numbers are universally used. Thisexplains why many OCR systems recognize numbers only, while relatively few understandthe full alphanumeric character range.

Second, there is similarity between many numerical and alphabetical symbol shapes. Forexample, while examining a string of characters combining letters and numbers, there isvery little visible difference between a capital letter "O" and the numeral "0." As humans,we can re-read the sentence or entire paragraph to help us determine the accurate meaning.

This procedure, however, is much more difficult for a machine. Third, we rely on contrast to help us recognize characters. We may find it very difficult to

read text which appears against a very dark background, or is printed over other words orgraphics. Again, programming a system to interpret only the relevant data and disregard therest is a difficult task for OCR engineers.

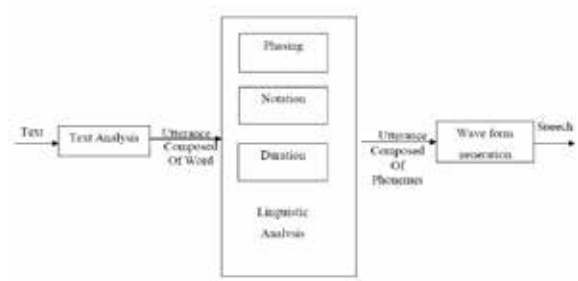


Fig 1: Text processing

2.RELATED WORK

A) For OCR

Shunji Mori, Ching Y. Suen et.al 1992 considered research and development of OCRsystems from U historical point of view. They divided their work into two parts: theresearch and development of OCR systems, and the historical development ofcommercial OCR's. The R&D part was further divided into two approaches: templatematching and structure analysis. It had been shown that both approaches are comingcloser and closer to each other and it seemed they tend to merge into one big stream openproblems are also raised in their work [4-7].

Jianli Liu, Nugent et.al 1993 developed a new AI-based OCR post processing technique,implemented as an intelligent OCR Editor (IOCRED), which could enable the automationof OCR post processing procedure and, therefore could result in the increase ofthroughput, the decrease of error rate and the reduction of cost per page of an OCRsystem. IOCRED was a novel AI approach to automating OCR post-processingprocedures. The IOCRED concept was based on the premise that different OCRalgorithms have distinct error characteristics and such distinction enable a cognitivedevice to automate its error detection and correction process. IOCRED system should beable to achieve a high throughput, low error rate and low cost OCR conversion. Theutilization of the IOCRED technique could result in the removal or the reduction of thecurrent OCR post processing techniques, error rate and cost per page [8].

Shaolei Feng and R. Manmatha 2006 have proposed a hierarchical, HMM- basedautomatic evaluation of OCR accuracy for digital library of books. They proposed aHidden Markov Model (HMM) based hierarchical alignment algorithm to align OCRoutput and the ground truth for books. They believed this was the first work toautomatical-

ly align a whole book without using any book structure information. Thealignment process worked by breaking up the problem of aligning two long sequencesinto the problem of aligning many smaller subsequences. This can be rapidly andeffectively done. Experimental results showed that their hierarchical alignment approachworks very well even if OCR output has a high recognition error rate. Finally, theyevaluate the performance of a commercial OCR engine over a large dataset of booksbased on the alignment results [1]

B) For Text to Speech

SoumyajitDey et.al 2007 proposed architectural optimizations for text to speechsynthesis in embedded system. The increasing processing power of embedded deviceshas created the scope for certain applications that could previously be executed indesktop environments only, to migrate into handheld platforms. An important feature ofthe computing systems of modern times is their support for applications that interact withthe user by synthesizing natural speech output. In their work, the performance of a Textto Speech Synthesis application is evaluated on embedded processor architectures andmedications in the underlying hardware platform are proposed for real time performanceimprovement of the concerned application [8].

3. METHODOLOGY

Figure 2 shows the framework of OCR. Most of the designs in OCR follow amodification of this architecture. Given a page for recognition, first it is preprocessed. Theaim of the preprocessing module is to prepare the image for recognition. Preprocessinginvolves binarization, skew correction and normalization. It undergoes some imageenhancements such as filtering out noise and increasing the contrast. Then, the image issegmented to separate the characters from each other. Segmentation occurs at two levels.On the first level, text, graphics and other parts are separated. On the second level, textlines, words and characters in the image are located. Information from connectedcomponent analysis and projection analysis can be used to assist text segmentation.Segmentation is followed by feature extraction, which is concerned with the representationof the object.

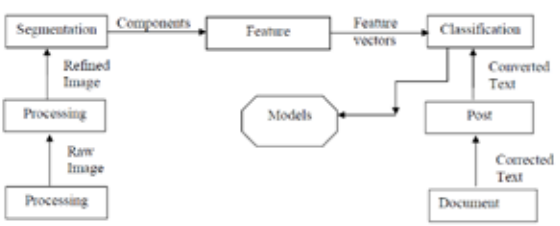
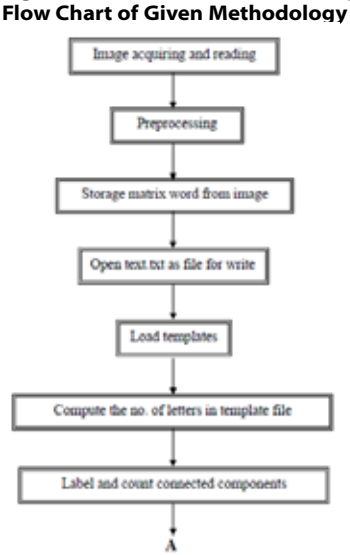


Fig. 2: Overview of conventional approach of OCR.

Flow Chart of Given Methodology



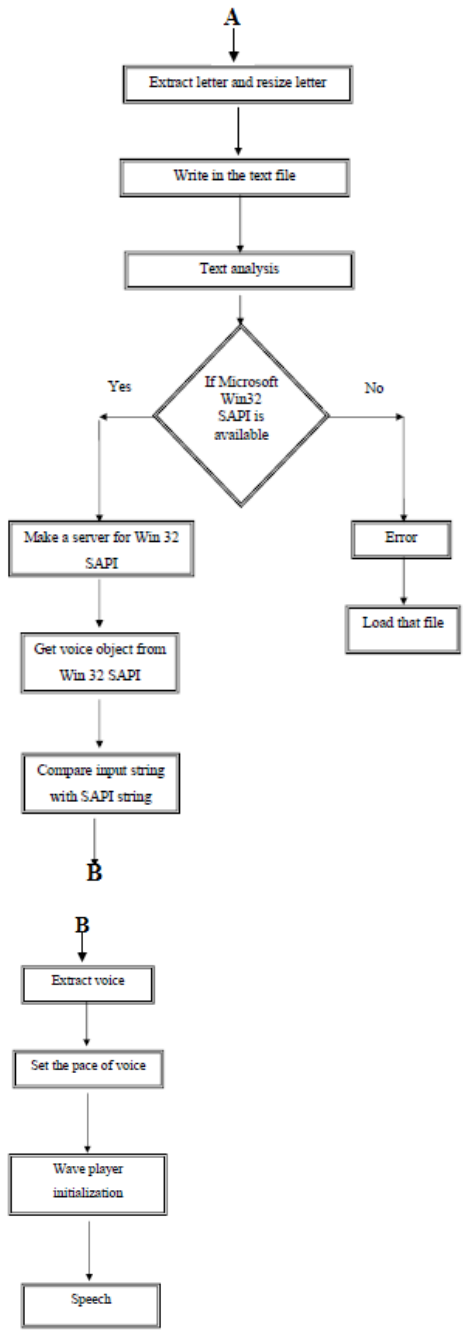


Fig. 3 Flow chart of used methodology.

4. RESULTS & DISCUSSION

In this work a two step program is made; in first step it gives the text output according toinput image with noise, then it convert that text into the speech, which is shown as**periodogram.**

Fig. 4 : Input Image

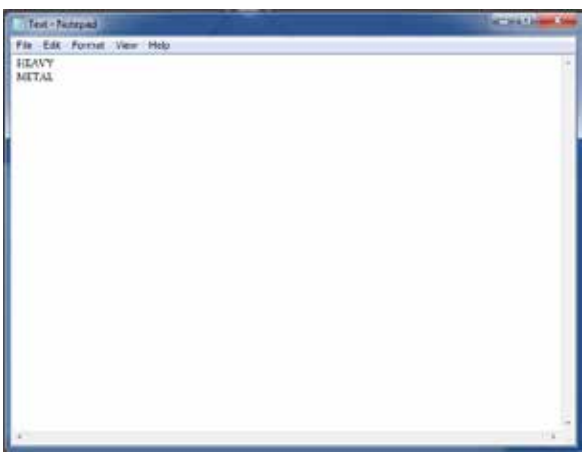


Fig.5 :Output text

we take different type of font character image in which font ofcharacter is different than previous and again it is converted into text and then speechsuccessfully.

5. CONCLUSION

Image into text and then that text into speech is converted by MATLAB. For image totext conversion firstly image is converted into gray image then black and white imageand then it is converted into text by MATLAB. Microsoft Win 32 SAPI library has beenused to build speech enabled applications, which retrieve the voice and audio outputinformation available for computer. This library allows selecting the voice and audiodevice one would like to use. By MATLAB we can select the voices from the list and canchange the pace and volume, which can be listen by installing wave player in theMATLAB. The application developed is user friendly, cost effective and applicable in thereal time.By this approach we can read text from a document, Web page or e-Book and cangenerate synthesized speech through a computer's speakers. This can save time byallowing the user to listen background materials while performing other tasks.

REFERENCES

[1] Ainsworth, W., "A system for converting English text into speech," Audio and | Electroacoustics, IEEE Transactions on , vol.21, no.3, pp. 288-290, | [2] Fushikida, Katsunobu; Mitome, Yukio; Inoue, Yuji, "A Text to Speech Synthesizerfor the Personal Computer," Consumer Electronics, IEEE Transactions on ,vol.CE- | 28, no.3, pp.250-256, | [3] Hertz, S., "English text to speech conversion with delta," Acoustics, Speech, andSignal Processing, IEEE International Conference on ICASSP '86. , vol.11, no., pp. | 2427-2430. | [4] Lynch, M.R.; Rayner, P.J., "Optical character recognition using a new connectionistmodel," Image Processing and its Applications, 1989., Third International | Conference on , vol., no., pp.63-67. | [5] Malyan, R.R.; Sunthankar, S.; Teranchi, H.; Yeghiazarian, A., "Perception of multiauthorhandprinted text," Character Recognition and Applications, IEE Colloquiumon , vol., no., pp.8/1-8/3. | [6] Curtis, K.M.; Race, P.; Aziz, A.A., "Parallelism and the transputer in the automatictranslation of text to speech," Acoustics, Speech, and Signal Processing.ICASSP-89., 1989 International Conference on , vol., no., pp.809-811 vol.2, 23-26 . | [7] Mori, S.; Suen, C.Y.; Yamamoto, K., "Historical review of OCR research anddevelopment," Proceedings of the IEEE , vol.80, no.7, pp.1029-1058. | [8] Jianli Liu; Nugent, J.H.; Bowen, D.G.; Bowen, J.E., "Intelligent OCR editor,"Electrical and Computer Engineering.