ISSN: 2278 – 1323
International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)
Volume 5, Issue 10, October 2016

SMART SHOPPING ASSISTANT LABEL READING SYSTEM WITH VOICE OUTPUT FOR BLIND USING RASPBERRY PI

Ms.Athira Panicker, **Ms.Anupama Pandey,** **Ms.Vrunal Patil**

YTIET,University of Mumbai YTIET,University of Mumbai YTIET,University of Mumbai

***Abstract*—This project aims at creating a system that reads text from product labels and help blind by working a s a shopping assistant. The system is controlled by raspberry pi model that has an arm processor. A camera-based assistive text reading algorithm is implemented to help the blind persons to read the text labels and product packaging from hand-held objects in their casual lives. Localization of text and recognition are conducted to acquire the text information from the area of interest. The localized text characters in the text regions are then binarized and recognized by optical character recognition in software .And then this text is feed into a text to speech synthesizer that will convert this printed text into voice and then this output will be heard by the blind person, thereby he will be able to read texts from any handheld object such as products during shopping.**

***Index Terms*— ocr technology ,tts synthesis ,raspberry pi,blind.**

I. INTRODUCTION

Visually impaired people report many difficulties with reading printed text using existing technology, including problems with alignment, focus, accuracy, mobility and efficiency. Reading is very important in today's world.Blind people are an integral part of our society. However, their disabilities have forced them to be dependent on others for assistance for daily life activities such as shopping,reading sign post etc.This has also made them to have lesser acess to computers,internet than the people with clear vision.Consequently, they have not been able to improve on their own knowledge, and have significant influence and impact on the society.

Today in the world there are more than 30 crore people who are visually impaired,out of which more than 4crore people are blind[1].According to the National Census of India there are around 2.2 crore disabled people in India,out of which more than 1.5 crore are blind.[2]This number tells us that the number of blind people are more than other disabled people in India.And this number is increasing rapidly through ages.

Recent advancement in computer vision, digital cameras, and portable computers made it easy to assist blind individuals by developing camera-based products using optical character recognition (OCR) systems.These optical readers can be used to recognize text information within the image,but these systems perform better with document images which contain text with simple background, well organized characters and standard fonts.These systems, thus, fail to recognize the product labels which complex colourfull backgrounds,multiple numbers of fonts and different font styles.Portable bar code reader helps a blind user to read from commercial products through speech and Braille[10].But a big limitation is that it is very hard for blind users to find the position of the bar code and to correctly point the bar code reader at the bar code.

If blind people or people with significant visual impairment can read from hand held objects,nearby sign posts or product labels then this will enhance their independent living and thereby foster economic and social self-sufficiency.So here we are going to propose a system that it useful to blind people to assist them to read text from these kinds of hand-held objects such as products while shopping.We have used a camera based assistive text reading framework to track the object of interest within the camera view and extract text information from the object and read it out to the blind person.Our proposed algorithm can effectively handle complex background and multiple patterns, and extract text information from both hand-held objects and nearby sign post.

II. LITERATURE SURVEY

A number of portable reading assistants are designed specifically for the visually impaired. "K-Reader Mobile" is a mobile application which allows the user to read mail, receipts, fliers, and many other documents [3].But these systems/device fail to give an economic solution of the problem and are available on specific platforms. No smart phones have designed for blind person until now. Thus accessibility of the Mobile application is a different question However, the document to be read must be nearly flat, placed on a clear, dark surface and contain mostly black text printed on white background and it does not reads from complex backgrounds.

Manuscript received Oct, 2016.

***Athira Panicker**, Electronics and Telecommunication,Mumbai University/ Yadavrao Tasgaonkar Institute Of Engineering and Technology Mumbai,India,*

***Anupama Pandey**, Electronics and Telecommunication,Mumbai University/ Yadavrao Tasgaonkar Institute Of Engineering and Technology.*

***Vrunal Patil**, Electronics and telecommunication,Mumbai University/ Yadavrao Tasgaonkar Institute Of Engineering and Technology.*

[4]Presents Darshan a Navigation System for blind people to navigate safely and quickly, in the system obstacle detection and recognition is done through ultrasonic sensors and USB camera. The proposed system detects the obstacles up to 300 cm via ultrasonic sensors and sends feedback in the form of beep sound via earphone to inform the person about the obstacle. USB webcam is connected with Raspberry Pi Embedded board which captures the image of the obstacle, which is used for finding the properties of the obstacle (Human Being). Human presence is identified with the help of human face detection algorithm written in Open CV. The constraints coming while running the algorithm on Embedded System are limited memory and processing time and speed to achieve the real time image processing requirements. The algorithm is implemented in Open CV, which runs on Debian based Linux environment

[5] Discusses an intelligent system visually impaired people experience difficulty and inconvenience using computers through a keyboard and mouse. This system provides a way to easily control many functions of a computer via speech. When a blind person speaks, the audio voice input is sent to the speech browser and them the output of the search is send through speakers. Many applications are running on this system but not all the applications able to fulfill the required needs but this system has better aspects in future for normal people as well as blind people. This application is firstly embedded on raspberry pi and Qt creator is the software which is being useful to interface this GUI with the hardware connected to Pi.

Wearable Obstacle Avoidance Electronic Travel Aids for Blind: A Survey last decades a variety of portable or wearable navigation systems have been developed to assist visually impaired people during navigation in known or unknown, indoor or outdoor environments. There are three main categories of these systems: electronic travel aids (ETAs), electronic orientation aids (EOAs), and position locator devices (PLDs). This paper presents a comparative survey among portable/wearable obstacle detection/avoidance systems (a subcategory of ETAs) in an effort to inform the research community and users about the capabilities of these systems and about the progress in assistive technology for visually impaired people. The survey is based on various features and performance parameters of the systems that classify them in categories, giving qualitative-quantitative measures. Finally, it offers a ranking, which will serve only as a reference point and not as a critique on these systems**.**

III. PROPOSED SYSTEM

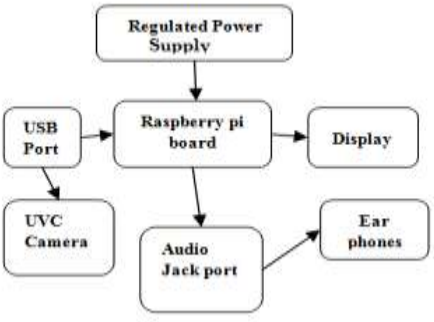

Fig 3.1 Basic block diagram of label reading

Our proposed algorithm can overcomes many of the drawbacks of the existing systems mentioned in the literature survey and effectively extracts text from hand-held objects and nearby sources which will be captured using the USB camera.A challenge that we face in this system is to position the label of the product or the text properly. In order to focus the object within the camera view a camera with wide angle is used as an approximate solution. Many a times unwanted text from the surrounding also gets included in the image.Thus to extract the required text from the image we proposed motion based method to isolate the region of interest and text recognition is done only on that region of interest. The algorithms developed for localization of text regions in the screen images are divided into two categories: rule based, learning based.. To solve the task at hand, extract text information from complex background with multiple and variable text pattern a text localization algorithm combine rule based and learning based algorithm which define novel features based on stroke orientation and edge distribution. These, intent, generate representative and discriminative text features which distinguish text character from backgrounds

FLOW OF WORKING

1. The blind person can have our system which is portable and easy to use. As his he/she has to carry the module with them for assistance.
2. In our project, the camera module is used to capture the real time image of the product. Which is the given to the main module.
3. The main module is of raspberry pi which is on its own a mini-computer, which processes the image captured by the camera.
4. Raspberry pi module, which contains the image processing code loaded, optical character recognition technique, is used to process the image. The image is processed internally in the raspberry pi hardware to separate the label from the captured image. by using OPENCV (open source computer vision) library. The desired letters in the label is identified by using Tesseract-OCR (optical character recognition).
5. Text To Speech Module: Text-to-Speech (TTS) Platform is a modular hardware design for text-to-speech applications. Platform is a fully integrated module that converts a stream of digital text into a high-quality English-speaking voice.The text is converted to voice and then is heard from the audio jack port using ear phones.

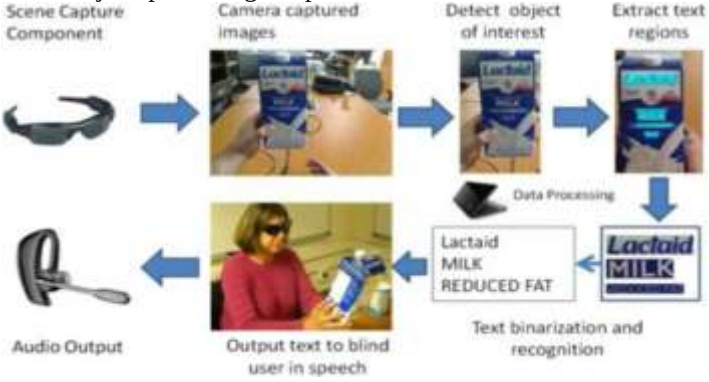

Fig 3.2 Flow of Project

Fig 3.3 Pictorial view of project

IV. HARDWARE

A. Raspberry Pi Model B+

Fig 4.1 Raspberry Pi model

The Raspberry Pi B+ has a Broadcom BCM2835 system on a chip (SoC), which includes an ARM1176JZF-S 700 MHz processor, Video Core IV GPU,with 512 MB RAM.It does not include a built-in hard disk or solid-state drive, but uses an SD card.On the model B the Ethernet port is provided by a built-in USB Ethernet adapter. The Raspberry Pi does not come with a real-time clock, so an OS must use a network time server, or ask the user for time information at boot time to get access to time and date for file time and date stamping..The ARM11 is based on version 6 of the ARM architecture (ARMv6k), which due to its age is no longer supported by several popular versions of linux.The Raspberry Pi also contains a 15-pin MIPI camera interface (CSI) connector, which is used with the Raspberry Pi Camera Add-on.

B. USB Camera

Input from camera is given to the raspberry pi module via USB cable.The USB camera we are using has Interpolated 8.0 Mega Pixel.High quality 5G wide angle lens.6 LEDs for night vision, with brightness controller.Snapshot button for still image capture.

C. Earphones

The processing output from the raspberry pi module is listened over here which is connected to audio jack of the raspberry pi This is used to get the proper audio output for the blind person.

D. LCD Display

Display we are using to show the demo to our project in an efficient way, so that we can able to see what the camera sees and what it is processing at a particular stage of operation.It can be connected to the raspberry pi model or computer via VGA, DVI, HDMI, Display Port

V. SOFTWARE

A. Operating System: Raspbian Os

Fig 5.1 Raspbian OS

Raspbian has a simple desktop environment with the familiar layout of a menu bar at the base of the screen and a programs menu in the bottom left corner. As Raspbian is a Linux operating system it has good security features, has excellent networking and capabilities and access to 1000's of free programs and utilities called packages that can be installed simply from the terminal.The Raspberry Pi is not a high powered device,so web browsers are a bit slower with less features.

B. Platform: Open Cv

Open CV (Open Source Computer Vision) is a library of programming functions for real time computer vision. Advantages of Open CV over MATLAB is that,Matlab is built on Java, and Java is built upon C. So when you run a Matlab program, your computer is busy trying to interpret all that Matlab code.Then it turns it into Java, and then finally executes the code. Open CV, on the other hand, is basically a library of

functions written in C/C++. You are closer to directly provide machine language code to the computer to get executed. So ultimately you get more image processing done for your computers processing cycles, and not more interpreting. As a result of this, programs written in Open CV run much faster than similar programs written in Mat lab.Hence we have used open cv as a platform here for image processing.

C. Language: Python

Python is a widely used high-level, general-purpose, interpreted dynamic programming language. Its design philosophy emphasizes code readability, and its syntax allows programmers to express concepts in fewer lines of code than would be possible in languages such as C++ or Java.

VI. OCR TECHNOLOGY

Optical character recognition (OCR) is the mechanical or electronic conversion of images of typed, handwritten or printed text into machine-encoded text. It is widely used as a form of data entry from printed paper data records, whether passport documents, invoices, bank statements, computerized receipts, business cards, mail, printouts of static-data, or any suitable documentation. It is a common method of digitizing printed texts so that it can be electronically edited, searched, stored more compactly, displayed on-line, and used in machine processes such as machine translation, text-to-speech, key data and text mining.

A. Pre-processing

OCR software often "pre-processes" images to improve the chances of successful recognition. Techniques include:

De-skew – If the document was not aligned properly when scanned, it may need to be tilted a few degrees clockwise or counterclockwise in order to make lines of text perfectly horizontal or vertical.

Despeckle – remove positive and negative spots, smoothing edges

Binarisation – Convert an image from color or greyscale to black-and-white. The task of binarisation is performed as a simple way of separating the text from the background[6].

Line removal – Cleans up non-glyph boxes and lines

Layout analysis or "zoning" – Identifies columns, paragraphs, captions, etc. as distinct blocks.Especially important in multi-column layouts and tables.

Line and word detection – Establishes baseline for word and character shapes, separates words if necessary.

Script recognition – In multilingual documents, the script may change at the level of the words and hence, identification of the script is necessary, before the right OCR can be invoked to handle the specific script.

Character isolation or "segmentation" – For per-character OCR, multiple characters that are connected due to image artifacts must be separated; single characters that are broken into multiple pieces due to artifacts must be connected.

Normalize aspect ratio and scale-Segmentation of fixed-pitch fonts is accomplished relatively simply by aligning the image to a uniform grid based on where vertical grid lines will least often intersect black areas. For proportional fonts, more sophisticated techniques are needed because whitespace between letters can sometimes be greater than that between words, and vertical lines can intersect more than one character.

B. Character recognition

There are two basic types of core OCR algorithm, which may produce a ranked list of candidate characters.

Matrix matching involves comparing an image to a stored glyph on a pixel-by-pixel basis; it is also known as "pattern matching", "pattern recognition", or "image correlation".This relies on the input glyph being correctly isolated from the rest of the image, and on the stored glyph being in a similar font and at the same scale. This technique works best with typewritten text and does not work well when new fonts are encountered. This is the technique the early physical photocell-based OCR implemented, rather directly.

Feature extraction decomposes glyphs into "features" like lines, closed loops, line direction, and line intersections. These are compared with an abstract vector-like representation of a character, which might reduce to one or more glyph prototypes. General techniques of feature detection in computer vision are applicable to this type of OCR, which is commonly seen in "intelligent" handwriting recognition and indeed most modern OCR software. Nearest neighbor classifiers such as the k-nearest neighbors algorithm are used to compare image features with stored glyph features and choose the nearest match.

C. Post-processing

OCR accuracy can be increased if the output is constrained by a lexicon – a list of words that are allowed to occur in a document.This might be, for example, all the words in the English language, or a more technical lexicon for a specific field. This technique can be problematic if the document contains words not in the lexicon, like proper nouns. Tesseract uses its dictionary to influence the character segmentation step, for improved accuracy.

The output stream may be a plain text stream or file of characters, but more sophisticated OCR systems can preserve the original layout of the page and produce, for example, an annotated PDF that includes both the original image of the page and a searchable textual representation.

D. Character detection

We use OpenCV (open source computer vision) library to process the images so that features for each letter could be extracted. First, we get the frames continuously from the camera and send it to the process. Once the object of interest is extracted

from the camera image using cascade classifier, subsequent process can be done using following steps.

Conversion to gray scale: The first stage in OCR involves generating a black-and-white version of the color or grayscale scanned page.OCR is essentially a binary process: it recognizes things that are either there or not. If the original scanned image is perfect, any black it contains will be part of a character that needs to be recognized while any white will be part of the background. Reducing the image to black and white is hence the first stage in figuring out the text that is to be processed.This conversion may also introduce some errors.

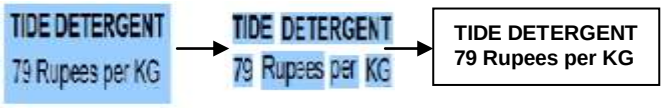

Fig 6.1 Conversion to gray scale

OCR: All OCR programs are slightly different, but generally they process the image of each page by recognizing the text character by character, word by word, and line by line

Basic error correction: OCR programs have an error checking features to help you spot mistakes.For example,it uses a method called near-neighbor analysis to find words that are likely to occur nearby, so text incorrectly recognized as "the barking bog" might be automatically changed to "the barking dog".

Layout analysis: Good OCR programs automatically detect complex page layouts, such as multiple columns of text, tables, images, and so on. Images are automatically turned into graphics, tables are turned into tables, and columns are split up correctly ,so the text from the first line of the first column isn't automatically joined to the text from the first line of the second column.

Proofreading: Even the best OCR programs aren't that perfect, especially when they're working from very old documents or poor quality printed text. That's why the final stage in OCR should always be a good, old-fashioned human proofread.

E. TESSERACT

Tesseract is Open source OCR engine. Tesseract works with independently developed Page Layout Analysis Technology. Hence Tesseract accepts input image as a binary image. Tesseract can handle both, the traditional- Black on White text and also inverse-White on Black text. Outlines of component are stored on connected Component Analysis. Nesting of outlines is done which gathers the outlines together to form a Blob. Such Blobs are organized into text lines. Text lines are analyzed for fixed pitch and proportional text. Then the lines are broken into words by analysis according to the character spacing. Fixed pitch is chopped in character cells and proportional text is broken into words by definite spaces and fussy spaces.

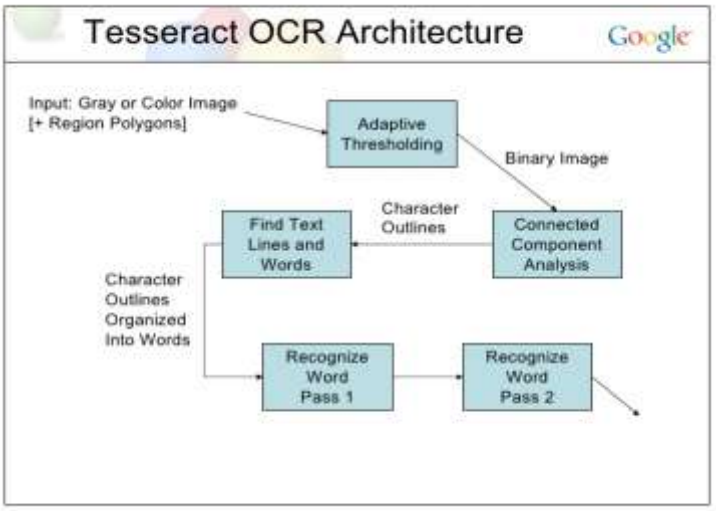

Fig 6.2 Tesseract Architecture.

Tesseract performs activity to recognize words. This recognition activity mainly consists of two passes. The first pass tries to recognize the words. Then satisfactory word is passed to Adaptive Classifier as training data, which recognizes the text more accurately. During second pass, the words which were not recognized well in first pass are recognized again through run over the pageFinally Tesseract resolves fuzzy spaces. To locate small and capital text Tesseract checks alternative hypothesis for x-height.

VII. TEXT TO SPEECH

A Text-To-Speech (TTS) synthesizer is a computer-based algorithm that can read any text aloud, whether it was directly introduced in the computer by an operator or scanned and submitted to an Optical Character Recognition (OCR) system. There is a fundamental difference between the system we are about to discuss here and any other talking machine (as a cassette-player for example) in the sense that we are interested in the automatic production of new sentences. Systems that just concatenate isolated words or parts of sentences, denoted as Voice Response Systems, are only applicable when a limited vocabulary is required, and when the sentences to be pronounced have a very restricted structure. In the case of TTS synthesis, it is impossible to record and store all the words of the language. It is thus more suitable to define Text-To-Speech as the automatic production of speech, through a grapheme-to-phoneme transcription of the sentences to utter.

TTS can be a help to an handicapped person who is unable to speak.With the help of an especially designed keyboard and a fast sentence assembling program, synthetic speech can be produced in a few seconds. Astro-physician Stephen Hawking gives all his lectures in this way.Blind people also widely benefit from TTS systems, when coupled with Optical Recognition Systems (OCR), which give them access to written information. The market for speech synthesis for blind users of personal computers will soon be invaded by mass-market

synthesizers bundled with sound cards. DEC talk (TM) is already available with the latest SoundBlaster (TM) cards now, although not yet in a form useful for blind people.

A. ARCHITECTURE OF TTS

The TTS system comprises of these 5 fundamental components: The input text is passed through these phases to obtain the speech.

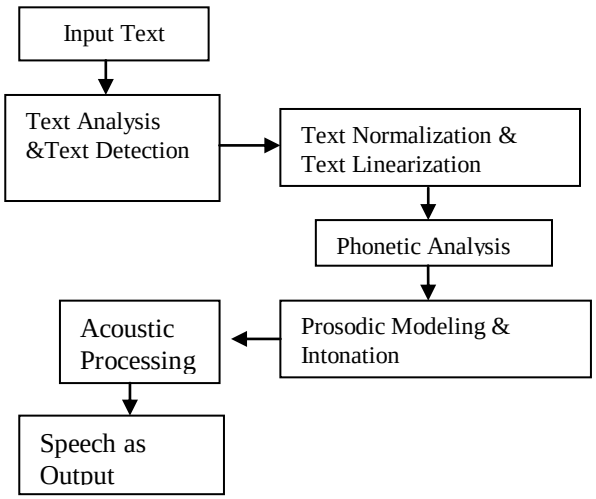

Fig 7.1 Block diagram of TTS

A. TEXT ANALYSIS AND TEXT DETECTION:

The Text Analysis part is the preprocessing part which analyzes the input text and organizes it into manageable list of words. It transforms numbers, abbreviations, acronyms and idiomatic into full text when needed. An important problem is found as soon as the character level: that of punctuation ambiguity (sentence end detection). It can be solved, to some extent, with elementary regular grammars [7].

Text detection is localizing the text areas from any kind of printed documents. Most of the previous researches were concentrated on extracting text from video. We aim at developing a technique that work for all kind of documents like newspapers, books etc.

B.TEXT NORMALIZATION AND TEXT LINEARIZATION

Text Normalization is the transformation of text to pronounceable form. Text normalization is often performed before text is processed in some way, such as generating synthesized speech or automated language translation. The main objective of this process is to identify punctuation marks and pauses between words. Usually the text normalization process is done for converting all letters of lowercase or upper case, to remove punctuations, accent marks , stop words or "too common words "and other diacritics from letters.

Text normalization is useful for example for comparing two sequences of characters which represented differently but mean the same. "Don't" vs. "Do not", "I'm" vs. "I am", "Can't" vs. "cannot" are some of the examples. The main 4 phases of Text Normalization are

(i). Number converter:

Number is pronounced differently in different situations. Like, 1772 (date): seventeen seventy two. 1772(phone number): one seven seven two, 1772 (quantifier): one thousand seven hundred and seventy two. Fractional and decimal numbers are handled. 0.302 (number): point three knot two.

(ii). Abbreviation converter:

Abbreviations area changed to full textual format. Mrs. - Misses St. Joseph St. - Saint Joseph Street.

(iii). Acronym converter:

Acronyms are replaced by single letter components. S. I. - S I.

(iv). Word segmentation:

Sentences are a group of word segments. Special delimiter to separate segments. Segments can be an acronym, a single word or a numeral.

Examples of acronyms:

"HIV" - "aitch eye ve"

"Henry IV" - "Henry the fourth"

"Chapter IV"- "Chapter four"

Punctuation marks are also identified.

Linearization is the process of giving a hyper text link to give the user a quick overview of the page. Then the TTS system will help to read out the linearized data. This feature helps in selecting the text and reading and also to list the links in the hyper text.

C. PHONETIC ANALYSIS:

Phonetic Analysis converts the orthographical symbols into phonological ones using a phonetic alphabet. Basically known as "grapheme-to-phoneme" conversion. Phone is a sound that has definite shape as a sound wave. Phone is the smallest sound unit. A collection of phones that constitute minimal distinctive phonetic units are called Phoneme. Number of phonemes is relatively smaller than the graphemes, only 44.

Phoneme Set (English)

- Vowels (19) : /a/, /ae/, /air/, /ar/, /e/, /ee/, /i/, /ie/, /o/, /oe/, /oi/, /oo/, /ow/, /or/, /u/, /ur/, /ue/, /uh/, /w/.
- Consonants (25) : /b/, /ks/gz/, /c/k/, /ch/, /d/, /f/, /g/, /h/, /j/, /l/, /m/, /n/, /ng/, /p/, /kw/, /r/, /s/, /sh/, /t/, /th/, /th/, /v/, /y/, /z/, /zh/.

Examples:

- /air/ : square, bear.
- /ow/ : down, house.
- /ks/gz/ : box, exist

Pronunciation of word based on its spelling has two approaches to do speech synthesis namely:

- Dictionary based approach

- Rule based approach.

A dictionary is kept were it stores all kinds of words with their correct pronunciation, it's just a matter of looking in to dictionary for each word for spelling out with correct pronunciation. This approach is very quick and accurate and the pronunciation quality will be better but the major drawback is that it needs a large database to store all words and the system will stop if a word is not found in the dictionary.

The letter sounds for a word are blended together to form a pronunciation based on some rule. Here main advantage is that it requires no database and it works on any type of input. same way the complexity grows for irregular inputs.

D. PROSODIC MODELLING AND INTONATION

The concept of prosody is the combination of stress pattern , rhythm and intonation in a speech. The prosodic modeling describes the speakers emotion. Recent investigations suggest the identification of the vocal features which signal emotional content may help to create a very natural synthesized speech.

Intonation is simply a variation of speech while speaking. All languages use pitch, as intonation to convey an instance, to express happiness, to raise a question etc. Rising Intonation (when the pitch of the voice increases)

i. Falling Intonation (when pitch of the voice decreases)
ii. Dipping Intonation (when the pitch of the voice falls and then rises)
iii. Peaking Intonation (when the pitch of the voice raises and then falls)

E. ACOUSTIC PROCESSING:

The speech will be spoken according to the voice characteristics of a person, There are three type of Acoustic synthesing available.

i. Concatenative Synthesis
ii. Formant Synthesis
iii. Articulatory Synthesis

The concatenation of prerecorded human voice is called Concatenative synthesis, in this process a database is needed having all the prerecorded words .The natural sounding speech is the main advantage and the main drawback is the using and developing of large database. Formant-synthesized speech can be constantly intelligible .It does not have any database of speech samples. So the speech is artificial and robotic. Speech organs are called Articulators. In this articulatory synthesis techniques for synthesizing speech based on models of the human vocal tract are to be developed. It produces a complete synthetic output, typically based on mathematical models

VIII. FUTURE SCOPE

This project can effectively distinguish the object of interest from the background or other objects in the camera view. In future this project can be implemented in hardware. It is used to detect and recognize object and vehicles on the road. So that it will assist person not to cross the road during vehicle movement. We will also extend our algorithm to handle non horizontal text strings. Our future work will extend our localization algorithm to process text strings with characters fewer than three and to design more robust block patterns for text feature extraction. We can adjust the alignment of camera and use more function of ocr to enhance it application .By enhancing application we can scan or process electronics labels, vehicles number and can be used for traffic monitoring.

IX. CONCLUSION

In many ways the result of the project are both surprisingly good and surprisingly bad. For images without definite edges the program may not work properly. But it will work perfectly for image texts which have prominent edge. For the product with fancy font, transparent text, text that is too small, blurred text, and for non planar surface it will not work properly. The labeling algorithm needs to be improved. A better labeling method of components could improve the detection of characters. This could get better results for circular text, which tends to be dismissed as the noise due to the grouping of the letters.

REFERENCES

[1] "10 facts about blindness and visual impairment", World Health Organization: Blindness and visual impairment, 2009. www.who.int/features/factfiles/blindness/blindness_facts/en/index.html
[2] National Census of India- http://censusindia.gov.in/Census_And_You/disabled_population.aspx
[3] KReader Mobile User Guide, knfb Reading Technology Inc. (2008). [Online]. Available: http://www.knfbReading.com
[4] L. Ma, C. Wang, B. Xiao, "Text detection in natural images based on multi-scale edge detection and classification," In the Int. Congress on Image and Signal Processing (CISP), 2010.
[5] "Intelligent system based on speech recognition with capability of self-learning" by Ms. Sneha K. Upadhyay, and Mr. Vijay N, International journal for research in applied science and engineering technology.
[6] *Zoran Zivkovic, "Improved Adaptive Gaussian Mixture Model for Background Subtraction", In Proc. ICPR, 2004*
[7] William A. Ainsworth, "A system for converting English text into speech", IEEE transactions on audio and electroacoustic, Vol. 21, No. 3, 1973
[8] ScanTalker, Bar code scanning application to help Blind Identify over one million products. (2006). [Online]. Available:http://www.freedomscientific.com/fs_news/PressRoom/en/2006/ScanTalker2-Announcement_3–30–2006.asp
[9] X. Chen, J. Yang, J. Zhang, and A. Waibel, "Automatic detection and recognition of signs from natural scenes," IEEE Trans. Image Process.,vol. 13, no. 1, pp. 87–99, Jan. 2004.