International Research Journal of Engineering and Management Studies (IRJEMS)
Volume: 03 Issue: 03 | March -2019 e-ISSN: 2395-0126

OPTICAL CHARACTER RECOGNITION FOR VISUALLY IMPAIRED PEOPLE IN SHOPPING TROLLEY

V.Balaji Ram, T.A.Bharathwaj,
K.DharunPranesh, P.KasirajaSathish,
Department of EIE,
Valliammai Engineering College,
SRM Nagar, Kattankulathur,
Tamilnadu, India

Dr Andy Srinivasan (Ph.D)
Professor,
Department of EIE,
Valliammai Engineering College,
SRM Nagar,Kattangulathur
Tamilnadu, India

E-Mail: vbalaji131@gmail.com, ta.bharathwaj98@gmail.com,
dharunpranesh28@gmail.com, kasirajasathish@gmail.com

Abstract: An OCR (Optical Character Recognition) system which is a branch of computer vision and in turn a sub-class of Artificial Intelligence. Optical character recognition is the translation of optically scanned bitmaps of printed or hand-written text into audio output by using of Raspberry pi. OCRs developed for many world languages are already under efficient use. This method extracts moving object region by a mixture-of-Gaussians-based background subtraction method. A text localization and recognition are conducted to acquire text information. To automatically localize the text regions from the object, a text localization and Tesseract algorithm by learning gradient features of stroke orientations and distributions of edge pixels in an Ada boost model. Text characters in the localized text regions are then binaries and recognized by off-the-shelf optical character recognition software. The recognized text codes are output to blind users in speech. Performance of the proposed text localization algorithm. As the recognition process is completed, the character codes in the text file are processed using Raspberry pi device on which recognize character using Tesseract algorithm and python programming, the audio output is listed.

I. INTRODUCTION

The Future Internet goal is to provide an infrastructure to have an immediate access to information about the physical world and its objects. Physical objects can be applicable to different application domains, such as e-health, warehouse management, etc. Each application domain may have different types of physical devices.Each physical device can have its own specifications, which is required to use in order to interact with it. To achieve the future Internet goal, a layered vision is

required that can facilitate data access. Things (IoT) is a vision that aims to integrate the virtual world of information to the real world of devices through a layered architecture.

As of 2016, the vision of the Internet of things has evolved due to a convergence of multiple technologies, including ubiquitous wireless communication, real-time analytics, machine learning, commodity sensors, and embedded systems. This means that the traditional fields of embedded systems, wireless sensor networks, control systems, automation (including home and building automation), and others all contribute to enabling the Internet of things.

The concept of a network of smart devices was discussed as early as 1982, with a modified Coke machine at Carnegie Mellon University becoming the first Internet-connected appliance, able to report its inventory and whether newly loaded drinks were cold. Mark Weiser's seminal 1991 paper on ubiquitous computing, "The Computer of the 21st Century", as well as academic venues such as UbiComp and PerCom produced the contemporary vision of IoT.

II .RELATED WORK

E. Cardillo, V. Di Mattia, G. Manfredi, P. Russo, A. De Leo, A. Caddemi, G. Cerri "**An Electromagnetic Sensor Prototype to Assist Visually Impaired and Blind People in Autonomous Walking",** IEEE Journal,2017.

The feasibility of an electromagnetic sensor to assist the autonomous walking of visually impaired and blind users is demonstrated in this paper. It is known that people affected by visual diseases usually walk assisted by some supports, among which the white cane is the most common. Our idea consists in applying a microwave radar on the traditional white cane making aware the user about the presence of an obstacle in a wider and safer range. Compared to the already existing Electronic Travel Aids devices, the proposed system exhibits better performance, noise tolerance and reduced dimensions. In the following, the latest developments of this research activity are presented, with special concern for the miniaturization of circuit board and antennas. A laboratory prototype has been designed and realized and the first test results of obstacle detection are hereby shown to demonstrate the effectiveness of the system.

Mrs.Shilpa Reddy K, MounikaS.k,Pooja K , Sahana N ,**"Text to Speech for the Visually Impaired",** IRJCS Journal,2017.

Perusing is fundamental in day by day life for everybody. Outwardly debilitated people can read just by utilization of unique applications by them like Braille dialect. The disadvantage of this framework is that each item does not give the content in Braille. In this paper, they have proposed an assistive content perusing system to help outwardly impeded people to peruse writings from different questions in their day by day lives. At first, we catch the picture of the required, pre-handling is performed on it. Pre-handling incorporates steps like dark scale and binarization, question of intrigue acknowledgment. In the proposed framework, we are making the utilization of OTSU calculation to change over the dim scale picture into binarized one. The content districts from the caught picture are then separated and perceived by utilizing optical character acknowledgment programming (OCR). The principle calculation in OCR to be specific MODI is utilized here. This extricated content of different textual styles and sizes then can be perceived independently and afterward consolidated in a word giving its yield as sound utilizing Text-to-discourse utilizing the SAPI libraries.

MallapaD.Gurav, Shruti S. Salimath, Shruti B. Hatti, Vijayalaxmi I. Byakod, ShivaleelaKanade **"A Reading aid for the Blind People using OCR and OpenCV"**, IJSRET Journal 2017.

Optical character recognition (OCR) is the identification of printed characters using photoelectric devices and computer software. It coverts images of typed or printed text into machine encoded text from scanned document or from subtitle text superimposed on an image. In this research these images are converted into audio output. OCR is used in machine process such as cognitive computing, machine translation, text to speech, key data and text mining. It is mainly used in the field of research in Character recognition, Artificial intelligence and computer vision. In this research, as the recognition process is done using OCR the character code in text files are processed using Raspberry Pi device on which it recognizes character using tesseract algorithm and python programming and audio output is listened. To use OCR for pattern recognition to perform Document Image Analysis (DIA) we use information in grid format in virtual digital library's design and construction. This research mainly focuses on the OCR based automatic book reader for the visually impaired using Raspberry pi. Raspberry pi features a Broadcom System On a Chip

(SOC) which includes ARM compatible CPU and an on chip graphics processing unit GPU. It promotes Python programming as main programming language.

III. BLOCK DIAGRAM

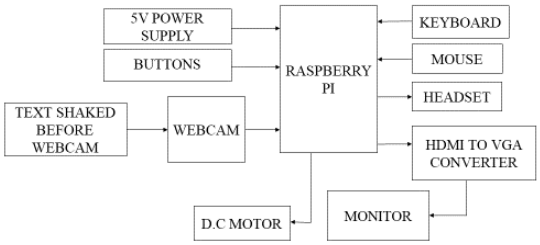

Fig 1: Block diagram of proposed system

The block diagram description is explained as follows-

- A 5V power supply is given to the Raspberry pi.
- There are four buttons of which three buttons which represents the food products. When a button is pressed representing the milk product, then the D.C Motor will make the trolley move towards that section.
- The Camera faces the product. While pressing the exit key at the keyboard, it will read the text of the product and produce an audio message by means of headphones and produce image for billing purpose
- Three buttons represents products and one button represents the signal needed for transmission to bill list. When we press the button, then the image of the product will be sent to the bill counter's mail and thus they will scan it by means of BAR CODE SCANNER.

IV.SYSTEM DESIGN

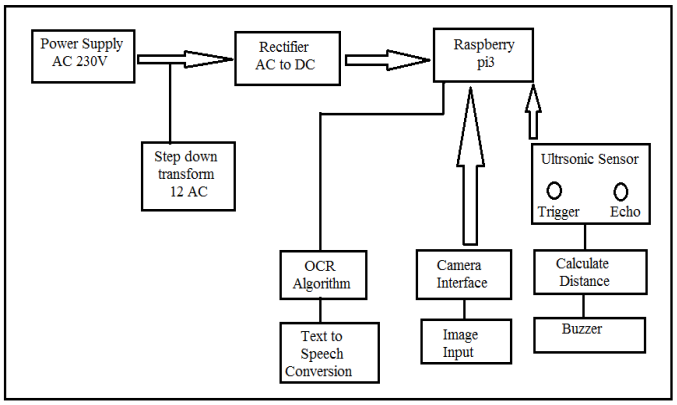

Fig 2. System Architecture

We have described a prototype system to read printed text on hand-held objects for assisting blind persons. In order to solve the common aiming problem for blind users, we have proposed a motion-based method to detect the object of interest, while the blind user simply shakes the object for a couple of seconds. The automatic ROI detection and text localization algorithms were independently evaluated as unit tests to ensure

effectiveness and robustness of the whole system. We subsequently evaluated this prototype system of assistive text reading using images of hand-held objects captured by ten blind users in person. Two calibrations were applied to prepare for the system test. First, we instructed blind users to place hand-held object within the camera view. Since it is difficult for blind users to aim their held objects, we employed a camera with a reasonably wide angle.

In future systems, we will add finger point detection and tracking to adaptively instruct blind users to aim the object. Second, in an applicable blind-assistive system, a text localization algorithm might prefer higher recall by sacrificing some precision.

When our application starts running it first check all the devices and resources which it needs are available or not. After that it checks the connection with the devices and gives control to the user. The GUI for the user has the following options. An optional label is used for displaying the image taken from the camera. A status box is for representing the detected data from the image. The capture button is to detect the data from the image. The detect button is to detect the human from the video streaming in front of the camera. The audio jack port is the output port here. The Raspberry board comes with integrated peripherals like USB, ADC and Serial etc. On this board we are installing Linux operating system with necessary drivers for all peripheral devices.

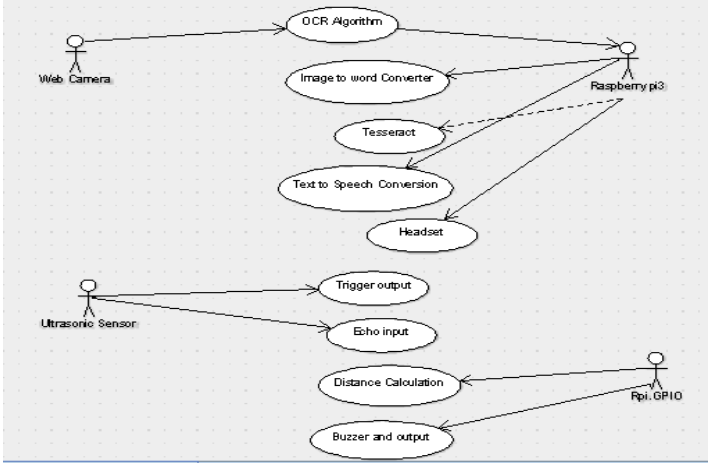

Fig 3.Use Case Diagram

When capture button is clicked this system captures the product image placed in front of the web camera which is connected to ARM microcontroller through USB .After selecting the process button the captured label image undergoes Optical Character Recognition(OCR) Technology. OCR technology allows the conversion of scanned images of printed text or symbols (such as a page from a book) into text or information that can be understood or edited using a computer program.

The most familiar example is the ability to scan a paper document into a computer where it can then be edited in popular word processors such as Microsoft Word.

However, there are many other uses for OCR technology, including as a component of larger systems which require recognition capability, such as the number plate recognition systems, or as tools involved in creating resources for SALT development from print based texts. In our system for OCR technology we are using TESSERACT library. Using Flite library the data will be converted to audio.

Camera acts as main vision in detecting the label image of the product or board then image is processed internally and separates label from image by using open CV library and finally identifies the product and identified product name is pronounced through voice.

Now it identifies received label image is converted to text by using tesseract library. Once the identified la-bel name is converted to text and converted text is displayed on display unit connected to controller. Now converted text should be converted to voice to hear label name as voice through ear phones connected to audio jack port using flite library.

Modules

- Image capturing and pre-processing.
- Automatic text extraction.
- Text recognition and audio output.

Image Capturing and Pre-Processing

The video is captured by using web-cam and the frames from the video is segregated and undergone to the pre-processing. First, get the objects continuously from the camera and adapted to process. Once the object of interest is extracted from the camera image and it converted into gray image. Use haar cascade classifier for recognizing the character from the object. The work with a cascade classifier includes two major stages: training and detection. For training need a set of samples. There are two types of samples: positive and negative.

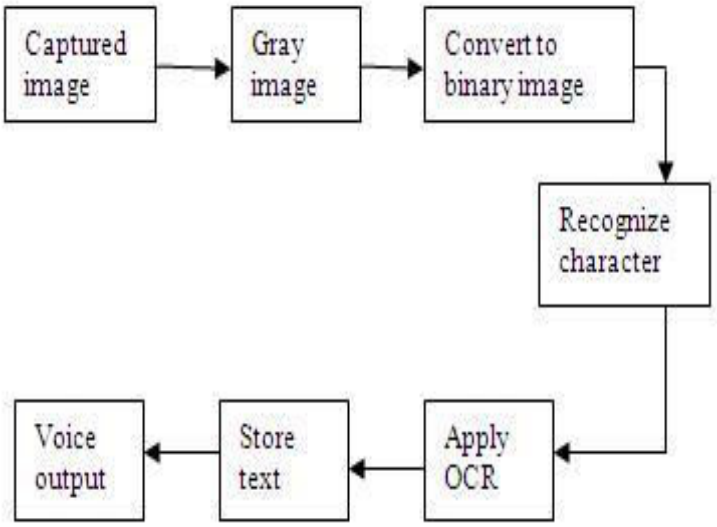

Fig.4. Image Capture and Pre-processing

To extract the hand-held object of interest from other objects in the camera view, ask users to shake the hand-held objects containing the text they wish to identify and then employ a motion-based method to localize objects from cluttered background.

Class Diagram

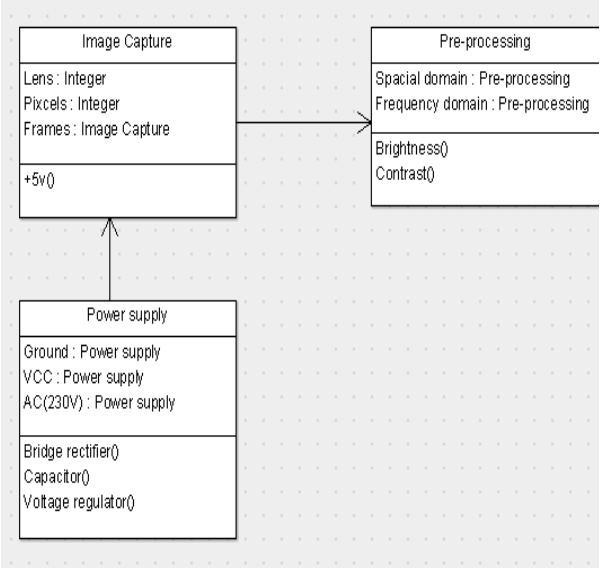

Fig. 5. Class Diagram for Image Capturing and Pre-Processing Module

In this module the image is capture by the web camera based on using lens, pixels, and frames by +5v(). The frame from the image is segregated and it undergoes to pre-processor with the spatial domain and frequency domain by brightness(), contrast(). Power supply is also used to capture image with the help VCC, AC(230) with function of bridge rectifier(),voltage regulator()**.**

Sequence Diagram

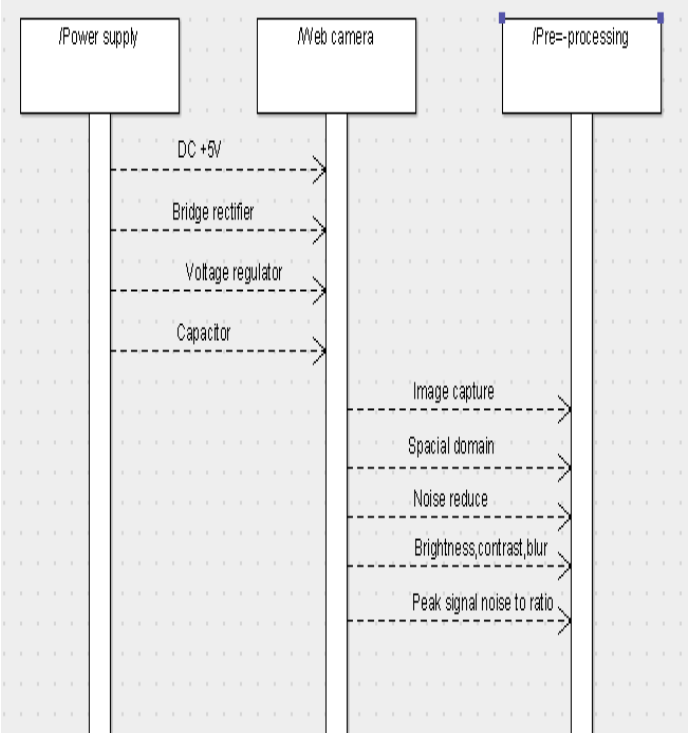

Fig. 6 Sequence Diagram for Image Capturing and Pre-Processing Module

In this module, the power supply flows direct current to the webcam by +5v,the bridge rectifier are used convert alternating current input to direct current output to the web camera, voltage are used to regulate one or more AC or DC voltages. Webcam are used to capture the segregated image from pre-processing. If the image captured by webcam has brightness, contrast, blur are stored by pre-processor. It also reduce the noise captured by webcam. Therefore it is a PNR. A sequence diagram is an interaction diagram that shows how objects operate with one another and in what order. A sequence diagram shows

how object interaction arranged in time sequence. It depicts the object and classes involved in the scenario and the sequence of message exchanged between the object needed to carry out the functionality of scenario.

Collaboration Diagram

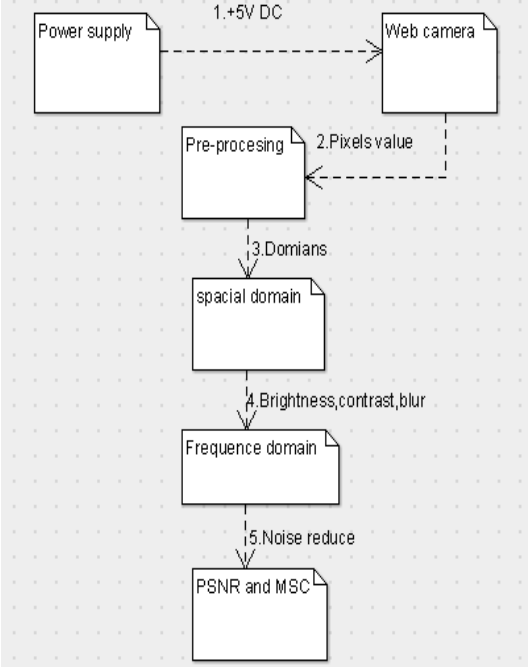

Fig.7.Collaboration Diagram for Image Capturing and Pre-Processing Module

In this module, the power supply uses +5vdirect current to the web camera and pixels value are segregated to the pre-processing that are pass through the special domain through domain it transform to frequency domain and checks the brightness, contrast, blur and reduce the noise through PSNR and MSC.

Automatic Text Extraction

In order to handle complex backgrounds, two novel feature maps to extracts text features based on stroke orientations and edge distributions, resspectively. Here, stroke is defined as a uniform region with bounded width and significant extent. These feature maps are combined to build an Adaboost based text classifier. The extraction information from audio and image source restricted to information execution from text. The actual transduction of audio and image data into text is the processing of OCR output.

Class Diagram

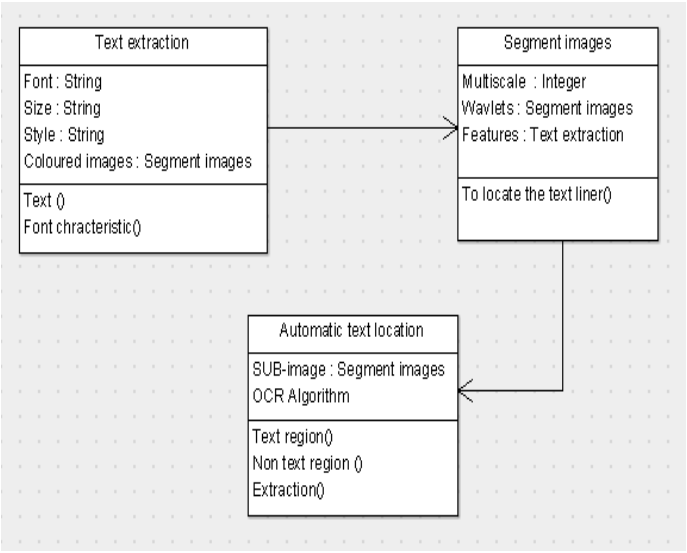

Fig. 7. Class Diagram for Automatic Text Extraction Module

In this module, the text extraction are extracted based on font, size, style,

coloured images by using text(),font characteristic().The extracted text are segmented into multi-scale, wavelets, features are used to locate the text. The automatic text location are recognised by SUB-image, OCR algorithm based on the text region(),non- text region(),extraction().

Sequence Diagram

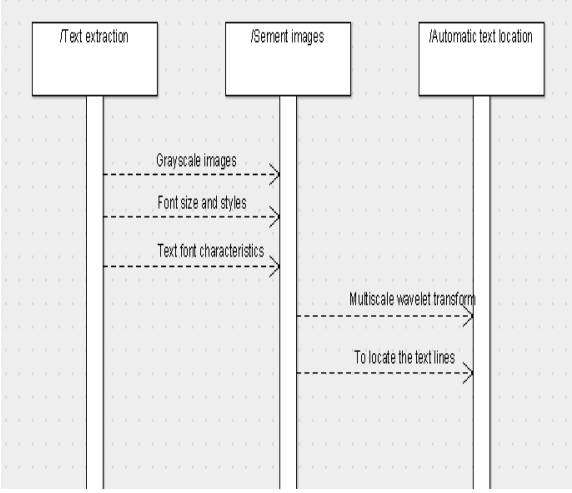

Fig. 8. Sequence Diagram for Automatic Text ExtractionModule

Text extraction are used to convert grey scale image into segment images by using font size, styles and text font characteristic. The segmented images are used to find the location of text are in the form of multistate wavelet transform. The segmented image are used to locate the text lines through the text region() ,non-text region(),extraction().

Collaboration Diagram

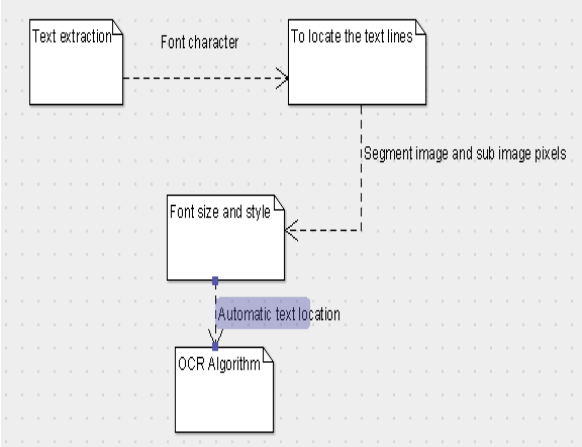

Fig 9 .Collaboration Diagram for Automatic Text Extraction Module

In this module, the text extraction are used to extract the text from the image by font character, text region and they are used to locate the text lines. The text lines are represented by front size and style through segmented image and sub pixels image. The font size and style are identified by OCR algorithm though automatic text location.

Text Recognition and Audio Output

Text recognition is performed by off-the-shelf OCR prior to output of informative words from the localized text regions. A text region labels the minimum rectangular area for the accommodation of characters inside it, so the border of the text region contacts the edge boundary of the text characters. However, this experiment show that OCR generates better performance text regions are first

assigned proper margin areas and binaries to segments text characters from background.

The recognized text codes are recorded in script files. Then, employ the Microsoft Speech Software Development Kit to load these files and display the audio output.

Blind users can adjust speech rate, volume and tone according to their preferences. Are designed to easily interface with dedicated computer systems by using the same USB technology that is found on most computers. Static random-access memory (SRAM) is a type of a semiconductor memory that uses bi-stable latching circuitry to store each bit.

Class Diagram

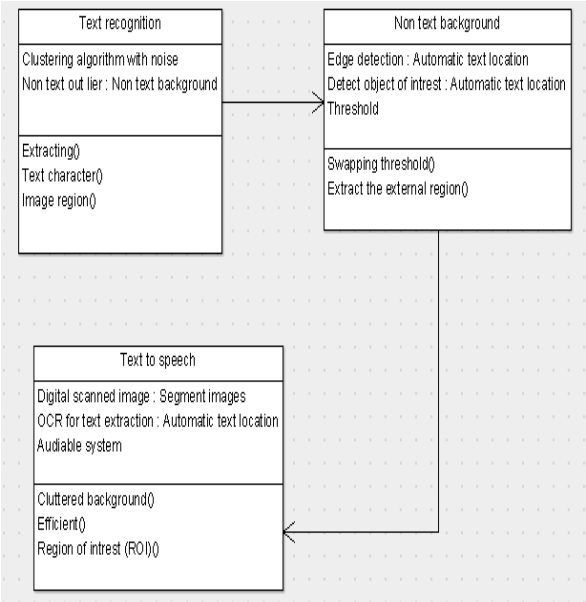

Fig.10 Class Diagram for Text Recognition and Audio Output Module

Description

In this module, the text recognition are used to recognize the text with the help of clustering algorithm non noise text outlier by text character().The non-text background are used to detect the object, threshold value and they are used to convert text to speech conversion module by using OCR algorithm to recognize the character of the text. Class diagram is a type of static structure of a system by showing the system classes, their attribute and the relationship between the classes. Private visibility hides information from anything outside the class partition. Public visibility allows all other to view the marked information. It is used for general conceptual modeling of the systematic of the application, and for detailed modeling translating the model into programming code. Class diagram can also used for data modeling. The classes in the class diagram represent both the main elements, interaction in the application, and the classes to be programmed.

Sequence Diagram

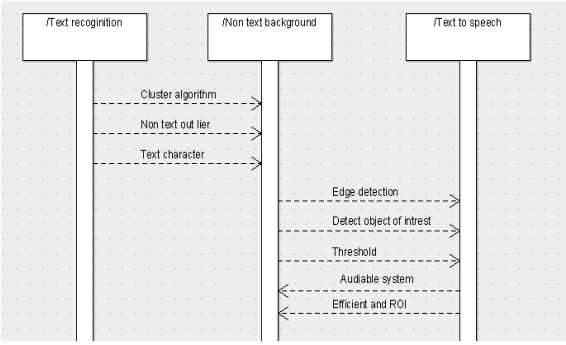

Fig.10. Sequence Diagram for Text Recognition and Audio Output Module

In this module, the text recognition are used to recognize the text by cluster algorithm with the non-text outlier to identify the text character for recognition. The non-text background are used to convert text to speech to detect the object, detection of edges, threshold values, the conversion are audible and efficient to the non- text background.

Collaboration Diagram

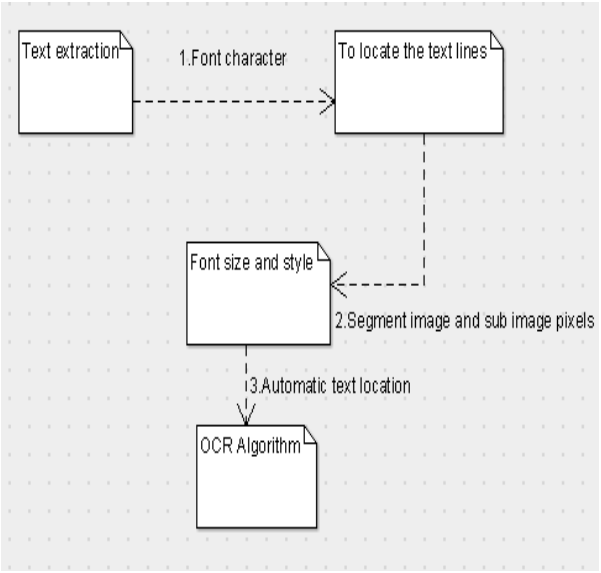

Fig. 11. Collaboration Diagram for Text Recognition and Audio Output Module

In this module, the text recognition are used to process through non text background and edge detection by non-text outlier. The non-textile are used to convert text to speech by threshold range. The conversion module are very efficient by digital scanned image. The images are passes through of Region Of Interest(ROI).

V.SOFTWARE RESULTS

The Software result is described as follows along with the diagram

- The figure on next page shows the software result of the Raspberry pi. The first four blue lines indicates that when a button is pressed regarding the desired food section, then it will move forward, turn right and then stop at the destination.
- Then press the escape key (ESC at keyboard) so that the camera will click the image and store it, and the raspberry pi will produce an audio message of the text in that image
- By pressing the fourth button, the image is sent to mail which is mentioned in programming.

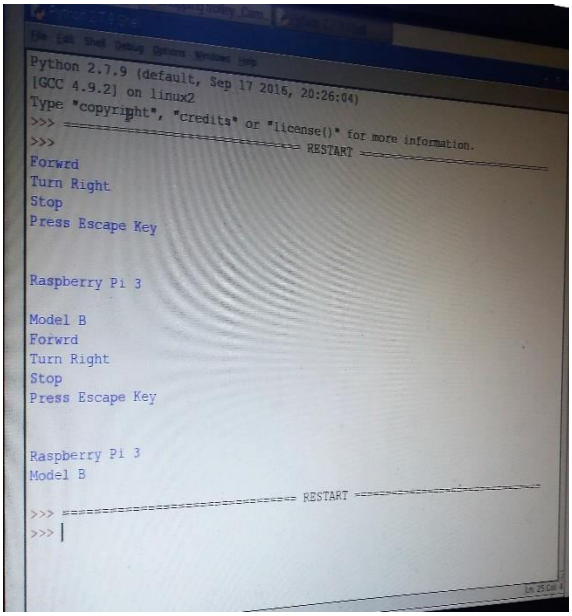

Fig 5: Software output

VI.CONCLUSION

An Optical Character Recognition (OCR) for visually impaired persons is proposed. When the buttons are pressed, the trolley moves to he required food sections and the product is captured on camera. Then after pressing the confirmation button, it is sent to mail for billing purpose.

REFERENCES

[1] A. Aladrn, G. Lpez-Nicols, L. Puig, and J. J. Guerrero, "Navigation assistance for the visually impaired using RGB-D sensor with range expansion," IEEE Syst. J., vol. 10, no. 3, pp. 922–932, Sep. 2016.

[2] T. Amemiya and H. Sugiyama, "Orienting kinesthetically: A haptic handheld wayfinder for people with visual impairments," ACM Trans. Accessible Comput., vol. 3, no. 2, 2010.

[3] B. And, S. Baglio, V. Marletta, and A. Valastro, "A haptic solution to assist visually impaired in mobility tasks," IEEE Trans. Human-Mach. Syst., vol. 45, no. 5, pp. 641–646, Oct. 2015.

[4] S. Bhatlawande, M. Mahadevappa, J. Mukherjee, M. Biswas, D. Das, and S. Gupta, "Design, development, and clinical evaluation of the electronic mobility cane for vision rehabilitation," IEEE Trans. Neural Syst. Rehabil. Eng., vol. 22, no. 6, pp. 1148–1159, 2014.

[5] B. B. Blasch, R. G. Long, and N. Griffin-Shirley, "National evaluation of electronic travel aids for blind and visually impaired individuals: Implications for design," in Proc. RESNA 12th Annu. Conf., 1989, pp. 133–134