Label Recognition with Audio Output for Blind People

Ruchika Dudhat[1] Bhagyashree Kakde[2] Sayali Gaikwad[3] Prof. Jyotsna Nanajkar[4]
[4]Assistant Professor
[1,2,3,4]Department of Electronics and Telecommunication Engineering
[1,2,3,4]Smt. Kashibai Navale College of Engineering, Pune-411041, India

Abstract— Visually impaired people face tremendous inconvenience in their routine life. The proposed system aims at making them self-sufficient by providing assistance to read the product labels thus fostering their daily activities. This work deals with an idea of making a portable hand-held device which recognizes the product labels and provides an audio output of the same. Portability of the device is achieved using Raspberry pi which has an inbuilt audio output jack. The camera contributes in capturing the product image which further undergoes text extraction. The extracted text is then processed under the OCR algorithm to recognize the text and then the audio output is generated.

Key words: Feature extraction, Text detection, Text-to-speech conversion, Optical character recognition, visually impaired people

I. INTRODUCTION

Of the 37 million people across the world who are blind, over 15 million are from India and this number is increasing rapidly. Blindness may result from a disease, injury or other conditions that limit vision. These people confront a number of challenges in their routine life. Blindness affects the person's basic abilities such as self navigating, crossing roads etc.

Technology has made numerous advances over recent years in order to help the disabled. The devices of this generation such as Smartphone, tablets and computers are equipped with various functions that provide an effortless access for the blind people. They provide functions such as Screen reading and audible navigation.

Fig. 1: Different products in the market.

Blind people also lack the ability to read; which is the crucial way of communication these days. And thus they have to rely on others for buying their basic amenities such as various food products etc. Different technologies such as bar code readers provide portability but are unable to recognize the product labels. It is also onerous job to point the barcode reader exactly towards the barcode. Also a text to speech system using optical character recognition is implemented on a laptop. But carrying a laptop also turns out to be a tedious job. Achieving both portability and appropriate label recognition is burdensome.

This work proposes a system which detects the product label, further recognizes it and then provides the output in the form of speech. Raspbian camera is used to capture the images due to its high resolution and compactness. The captured image is processed using the optical character recognition algorithm. Portability of the system is achieved using raspberry pi.

II. ARCHITECTURE

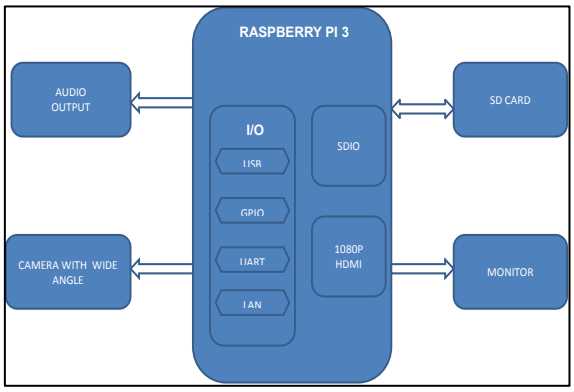

Fig. 2: Hardware Architecture

Portability being the prime essential Raspberry Pi has been taken into consideration due to its compactness and simplified user interface. The CSI camera slot is used to interface the camera which serves the basic requirement of the project. An 8 Gb memory card inserted in the SDIO port stores the OCR algorithm as well as the database of characters in both the cases. The HDMI port is used to establish a connection between the raspberry pie and the monitor. Some modern monitors are provided with HDMI port but some aren't. In that case a HDMI to VGA connector has to be used. Raspberry pi also has an inbuilt composite video and audio output jack that provides the converted text to speech output. The Raspberry pi camera module is a 5 Mp add-on for Raspberry pi featuring a fixed focus lens

III. WORKING OF SYSTEM

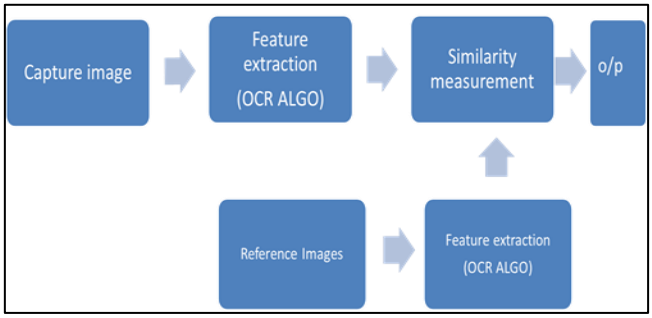

Fig. 3: Flow of the project

The alphabets in both the cases are stored as training images in the database. This database is made available on the 8Gb SD card which is to be inserted in the SDIO slot present on the Raspberry Pi. Feature extraction of these images is done using the OCR algorithm. The Raspberry Pi camera module is interfaced with Raspberry pi using the Csi camera slot.

The camera captures the product images for the further processing by means of OCR algorithm. The product image that is to be converted into text is loaded and then scanned. Further the image is converted into grayscale image. FIR filter is used to subtract any kind of flaws and imperfections in the image. The lines of text are extracted from the image, based on the white space. Further the text lines are sliced into characters, again based on the white space between the characters. These characters undergo feature extraction and then similarity measurement between the sliced characters and the training images takes place. For each character most closely matching character from the database is determined and the output text is appended.

The text is then accumulated and converted into speech and is taken out as audio output through the audio jack of Raspberry Pi with the help of a pair of earphones.

IV. RASPBERRY PI3 MODULE

The Raspberry Pi is a series credit card-sized single-board computers developed in United Kingdom. It has evolved through several versions that feature variations in memory capacity and peripheral-device support. Raspberry Pi 3 module is selected for the project. It is provided with a SD card slot as well as an inbuilt audio output jack. Camera interfacing is simplified by the CSI camera port thus satisfying all the requirements of the system.

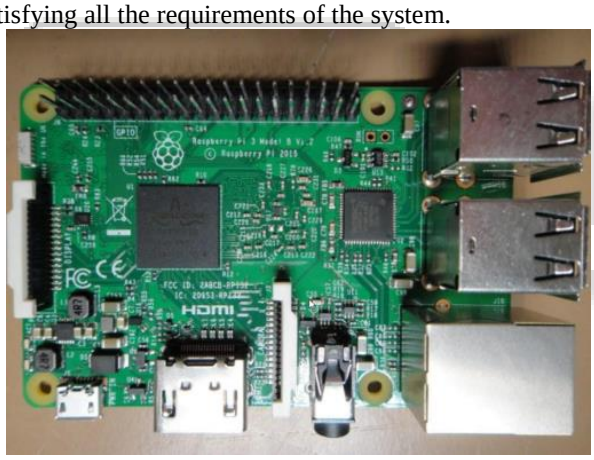

Fig. 4: Raspberry Pi 3 module

A. Raspberry Pi 3 specifications

- SoC: Broadcom BCM2837
- CPU: 1.2 GHZ quad-core ARM Cortex A53 (ARMv8 Instruction Set)
- GPU: Broadcom VideoCore IV @ 400 MHz.
- Memory: 1 GB LPDDR2-900 SDRAM.
- Power source: 5 V via MicroUSB or GPIO header

B. Features

- 15 Pin MIPI camera interface (CSI) connector, used with Raspberry Pi camera.
- Equipped with Bluetooth 4.1 protocol.
- Includes 4 USB ports.
- Supports video output through composite video jack.
- Supports audio output through composite audio jack.
- Equipped with 2.4 GHz 802.11n WiFi protocol.
- Support monitor and TV interfacing through HDMI port.

V. RASPBERRY PI CAMERA

The Raspberry Pi camera module is a 5 MP custom designed add on for Raspberry Pi, featuring fixed focus lens. It attaches to Pi by way of one of the small sockets on the board upper surface and uses the dedicated CSI interface, designed especially for interfacing to cameras.

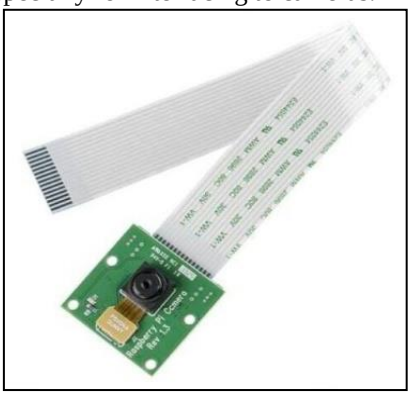

Fig. 5: Raspberry Pi camera module

A. Specifications

- Still resolution 5 Mp.
- Sensor omnivision OV5647.
- Sensor resolution 2592x1944 pixels.
- Focal ratio 2.9

B. Features

- Supports chief ray angle correction.
- Supports frame rate up to 120 fps.
- Supports format conversions available via GPU.
- Equipped with on chip PLL.

VI. ALGORITHM

The algorithm used to accomplish the text detection and recognition is termed as optical character recognition. OCR is the electronic conversion technique implemented to extract the text from either handwritten or printed text from an image and convert it into machine encoded text. The detailed flow of the algorithm is as follows:

- Load training images
- Load the scanned image of the document to be converted to text
- Convert the scanned image to grayscale
- Filter the scanned image using a low-pass Finite Impulse Response (FIR) filter to remove dust
- Break the document into lines of text, based on whitespace between the text lines
- Break each line into characters, based on whitespace between the characters; using the average character width, determine where spaces occur within the line
- For each character, determine the most closely matching character from the training images and append that to the output text; for each space, append a space character to the output text
- Output the accumulated text

VII. CONCLUSION

Considering the heavy growth in the number visually impaired individuals technology should excel and bring about a variety of inventions to reduce their struggle. The

project is implemented for the same reason. Reading being an efficient and crucial way of communication, the system is designed to assist them in this basic ability. The system captures images of the products, recognizes the labels and converts it into speech for the convenience of blind people. An approach is made to achieve portability and hence Raspberry Pi module is incorporated in the system. The system accomplishes all the tasks required to foster the living of blind people and to make them self-sufficient.

REFERENCES

[1] Kwang In Kim, Keechul Jung and Jin Hyung Kim, "Text Detection in Images Using Support Vector Machines and Continuously Adaptive Mean Shift Algorith", 2003.
[2] Xilin Chen, Jie Yang, Jing Zhang, and Alex Waibel, "Automatic Detection and Recognition of Signs from Natural Scenes", 2004.
[3] Sunil Kumar, Rajat Gupta, Nitin Khanna, Santanu Chaudhury and Shiv Dutt Joshi, "Text Extraction and Document Image Segmentation Using Matched Wavelets and MRF Model", 2007.
[4] Boris Epshtein, Eyal Ofek ,Yonatan Wexler,Microsoft corporation,"Detecting Text with Stroke Width Transform", 2010.
[5] K.Kalaivani, R.Praveena, V.Anjalipriya, R.Srimeenar, "Real time Implementation of Image Recognition", 2014.

