

Blind's Eye – Wearable Object Detection, Recognition and Identification for Visually Impaired

Ms. B Bhoomika Bopanna

Department of Computer Science
and Engineering

Srinivas Institute of Technology,
Valachil

Mr. Sudarshan K

Department of Computer Science
and Engineering

Srinivas Institute of Technology,
Valachil

Ms. Chandana K K

Department of Computer Science
and Engineering

Srinivas Institute of Technology,
Valachil

Abstract - In this generation printed text appears everywhere. Because of this blind people always take help of others to buy some product. Thus blind people need some assistance to read text information of the product. In worldwide there are over 314 million visually challenged people. Therefore, we come up with the system that reads text from product labels and help blind by working as a shopping aid. The system is developed on Raspberry Pi model using python programming language. It uses the Optical Character Recognition technology for the identification of the printed characters using image sensing devices and computer programming. The system captures the document image placed in front of the camera. The OCR (Optical Character Recognition) package installed in Raspberry Pi scans it into a digital document which is then subjected to grey scale conversion, erosion, image dilation, segmentation to perform character recognition. After recognition, the product is identified as soap, paste, book or as such. And then the recognized text is fed 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.

1. INTRODUCTION

Reading is obviously essential in today's society. Printed text is seen in every product packages. Visually impaired people cannot read these texts. The implementation of this paper help blind to read printed labels and product packages. This will enhance independent living, and faster economic and social sufficiency. Today there are already a few systems that have some promise for portable use, but they cannot handle product labelling. For example portable bar code readers designed to help blind people identify different products in an extensive product database can enable users who are blind to access information about these products through speech. 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 it at the bar code reader.

Blind's Eye is proposed to help blind persons read text labels and product packaging from hand-held objects in their daily lives. To identify the desired letters in the label from the camera image, Tesseract OCR is used. The recognized text codes are output to blind users in speech.

1.1 Problem Statement

Today, there are already a few systems that have some promise for portable use, but they cannot handle product labelling and consumes more time. Some reading assistive systems such as pen scanners might be employed in these and similar situations. Such systems integrate OCR software to offer the function of scanning and recognition of text and some have integrated voice output. Two of the biggest challenges to independence for blind individuals are

- Difficulties in accessing printed material.
- The stressors associated with safe and efficient navigation.

1.2 Existing System

Most of the OCR software cannot directly handle scene images with complex backgrounds. A number of portable reading assistants have been designed specifically for the visually impaired. Although a number of reading assistants have been designed specifically for the visually impaired, to our knowledge, no existing reading assistant can read text from the kind of challenging patterns and back-grounds found on many everyday commercial products. In assistive reading systems for blind persons, it is very challenging for users to position the object of interest within the centre of the camera's view. As of now, there are still no acceptable solutions.

1.3 Proposed System

The proposal presents a prototype system of assistive text reading. In this project, we have proposed a text read out system for the visually challenged. The proposed fully integrated system has a camera as an input device to feed the printed text document for digitization and the scanned document is processed by a software module, the OCR (optical character recognition engine). A methodology is implemented to the recognition sequence of characters and the line of reading. As part of the software

development, the Open CV (Open source Computer Vision) libraries are utilized to capture image of text and to do the character recognition. Optical character recognition (OCR) translates captured images of printed text into machine-encoded text. OCR is a process which associates a symbolic meaning with objects (letters, symbols an number) with the image of a character. Once, the character is recognized, it is made available as an audio output.

1.4 Objective

In solving the task at hand, to extract text information from complex backgrounds with multiple and variable text patterns, this project uses a camera to capture the image of the focused text, extract separate text characters from the scene and then provide a speech output using the Tesseract OCR and also provide the audio output of the same.

2. LITERATURE SURVEY

The paper as in [1] describes about the visually impaired to find their navigation as they often lack the needed information for by passing obstacles and hazards. Electronic Travel Aids (ETAs) are devices that use sensor technology to assist and improve the blind user’s mobility in terms of safety and speed.

The paper as in [2] says about the arrival of fast and cheap digital electronics. Sensory devices open new pathways to the development of sophisticated equipment to overcome limitations of the human senses. This paper addresses the technical feasibility of replacing human vision by human hearing through equipment that translate images into sounds.

The paper as in [3] says about the goal of Blind Aid project to develop navigational assistance technology for the blind or visually impaired. Specifically, it seeks to develop a portable Electronic Travel Aid (ETA) for visually impaired users, along with the accompanying radio frequency identification (RFID) localization infrastructure used to equip buildings.

Today, there are already a few systems that have some promise for portable use, but they cannot handle product labelling.

- Braille lippy bar code stickers sticks on portable hand held packaged product but it is very expensive. Because of this every packaged product which you use have to stick Braille lippy code stickers and it is not possible.[4]
- Portable bar code readers designed to help blind people identify different products in an extensive product database and can enable users who are blind to access information about these products through Speech. [5] 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.
- Pen Scanner for blind person have some drawbacks like optical character bar code reader where the blind person cannot exactly point it in front of the product.[6]

3. SYSTEM DESIGN

The purpose of the design phase is to plan a solution of the problem specified by the requirement document. The design of a system is perhaps the most critical factor affecting the quality of the software, and has a major impact on the later phases, particularly testing. The output of this phase is the design document.

3.1 High Level Design

High-level design which is sometimes also called system design, aims to identify the modules that should be in the system, the specifications of these modules, and how they interact with each other to produce the desired results. At the end of system design all the major file formats, output formats, as well as the major modules in the system and their specifications are decided.

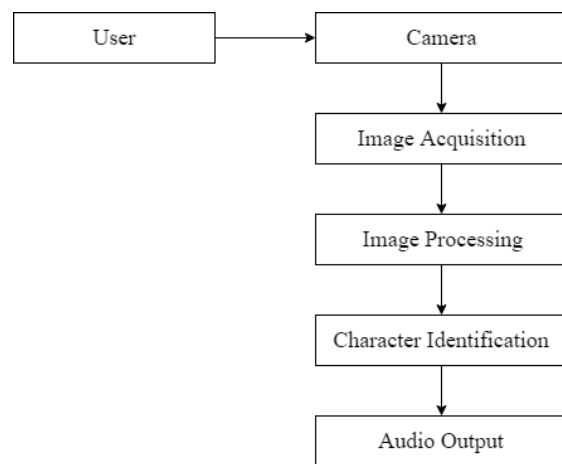


Figure 3.1: System Architecture

Figure 3.1 shows the system architecture of Blind’s Eye. User needs to select the product which will then be captured by the camera. The captured image is processed using Tesseract OCR. If the character is recognized, then the result is sent as an audio output to the user.

3.2 Detailed Design

During detailed design the internal logic of each of the modules specified in system design is decided. During this phase further details of the algorithmic design of each of the modules is specified. The logic of module is usually specified in a high-level design description language, which is independent of target language in which the software will eventually be implemented.

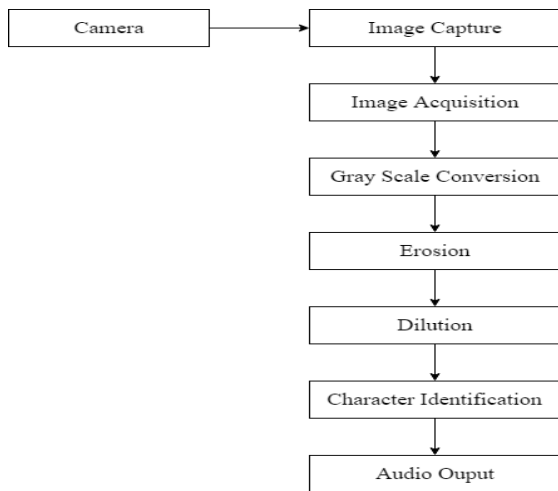


Figure 3.2: Data Flow Diagram

Figure 3.2 shows the data flow diagram for identification of product. The camera will capture the image. Captured image is converted into grey scale and image erosion and dilution are done. Further the character is recognized using tesseract OCR and result is sent to the user through speech.

4. SYSTEM IMPLEMENTATION

OpenCV-Python

OpenCV supports a wide variety of programming languages such as C++, Python, Java, etc., and is available on different platforms including Windows, Linux, OS X, Android, and iOS. OpenCV-Python is the Python API for OpenCV, combining the best qualities of the OpenCV C++ API and the Python language. It is a library of Python bindings designed to solve computer vision problems.

Raspberry Pi 3 Model B+

Raspberry pi is a portable, powerful and mini computer. The board length is only 85mm and width is only 56mm. Its size is as big as a credit card but it is capable of little PC. It can be used for many of the things that your desktop PC does, like high-definition video, spreadsheets, word-processing, games and more. Raspberry Pi also has more wide application range, such as music machines, parent

detectors to weather stations, tweeting birdhouses with infra-red cameras, lightweight web server, home automation server, etc. It enables people of all ages to explore computing, learn to program and understand how computers work. The Raspberry Pi Model B+ provides more GPIO, more USB than Model B. It also improves power consumption, audio circuit and SD card. It is more useful for embedded projects.



Figure 4.1: Raspberry Pi Board

Raspberry Pi Camera

The Raspberry Pi camera module can be used to take high-definition video, as well as still photographs. It’s easy to use for beginners. People use it for time-lapse, slow-motion and other video cleverness. Libraries can also be used with the camera to create effects. The module has a five megapixel fixed-focus camera that supports 1080p30, 720p60 and VGA90 video modes, as well as stills capture. It attaches via a 15cm ribbon cable to the CSI port on the Raspberry Pi. It can be accessed through the MMAL and V4L APIs, and there are numerous third-party libraries built for it, including the Picamera Python library.

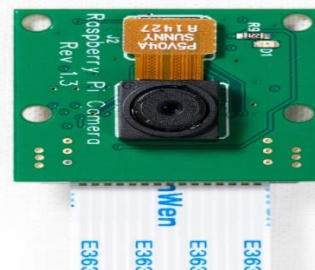
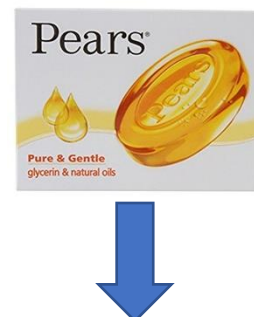


Figure 4.2: Raspberry Pi Camera

Flow of working

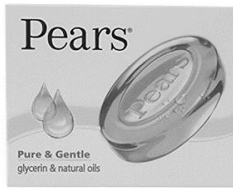


5. CONCLUSION

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. 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] G.Balakrishnan, G.Sainarayanan, "wearable stereo vision for blindly impaired people", *mechatronics & automation(ICMA) international conference*, 2012.
- [2] B.L.Meijer, "experimental system for auditory image representation", *IEEE transaction on biomedical engineering*, 1992.
- [3] F.Hong, A.Chekima, "ETA for visually impaired, international journal of disaster recovery and business connectivity", *IEEE a new vision vol.3,aug*, 2011
- [4] Richard L. Windsor, "Introduction to Braille language", http://www.lowvision.org/introduction_to_braille.html, 1995
- [5] Scan Talker, "Bar code scanning application to help Blind Identify over one million products", http://www.freedomscientific.com/fs_news/PressRoom/en/2006/ScanTalker2-Announcement_3-30-2006.asp, 2006
- [6] D. Sreenivasan, Dr. S. Poonguzhali, "An Electronic Aid for Visually Impaired in Reading Printed Text", *International Journal of Scientific & Engineering Research*, Volume 4, Issue 5, May-2013

*Erosion*

Erosion erodes away the boundaries of foreground object. It is used to diminish the features of an image.

Working of erosion

- A kernel (a matrix of odd size(3,5,7)) is convolved with the image.
- A pixel in the original image (either 1 or 0) will be considered 1 only if all the pixels under the kernel is 1, otherwise it is eroded (made to zero).
- Thus all the pixels near boundary will be discarded depending upon the size of kernel.
- So the thickness or size of the foreground object decreases or simply white region decreases in the image.

Dilation

Dilation increases the object area. It is used to accentuate features.

Working of dilation

- A kernel (a matrix of odd size(3,5,7)) is convolved with the image.
- A pixel element in the original image is '1' if atleast one pixel under the kernel is '1'.
- It increases the white region in the image or size of foreground object increases.