Indian Journal of Science and Technology, Vol 9(30), DOI: 10.17485/ijst/2016/v9i30/98716, August 2016
ISSN (Print) : 0974-6846
ISSN (Online) : 0974-5645

Camera based Text to Speech Conversion, Obstacle and Currency Detection for Blind Persons

D. B. K. Kamesh*, S. Nazma, J. K. R. Sastry and S. Venkateswarlu

Department of Electronics and Computer Engineering, KL University, Vaddeswaram - 522 502, Guntur District, Andhra Pradesh, India; kameshdbk@kluniversity.in, shaiknazu321@gmail.com, drsastry@kluniversity.in, somu23@kluniversity.in

Abstract

Background/Objectives: The main object of this paper is to present an innovated system that can help the blind for handling currency. **Methods/Statistical Analysis**: Many image processing techniques have been used to scan the currency, remove the noise, mark the region of interest and convert the image into text and then to sound which can be heard by the blind. The entire system is implemented by using Raspberry Pi Micro controller based system. In the proto type model an IPR sensor is used instead of camera for sensing the object. **Findings:** In this paper a novel method has been presented using which one can recognize the object, mark the interesting region within the object, scan the text and convert the scanned text into binary characters through optical recognition. A second method has been presented using which the noise present in the scanned image is eliminated before characters are recognized. A third method that can be used to convert the recognised characters into e-speech through pattern matching has also be presented. **Applications**: An embedded system has been developed based on ARM technology which helps the blind persons to read the currency notes. All the methods presented in this paper have been implemented within an embedded application. The embedded board has been tested with different currency notes and the speech in English has been generated that identify the value of the currency. Further work can be done to generate the speech in different other both National and International Languages.

Keywords: Camera based Detection for Blind Persons, Currency Detection, Raspberry Pi Board, Text to Speech Conversion

1. Introduction

A camera based reader helps blind persons to read labels on the products and other handheld devices in their day by day lives[1]. To differentiate the object from heavy backgrounds and other surroundings, an effective motion based method is used to define Region of Interest (ROI) in the camera view. In the obtained ROI text, localization and recognition are done[2]. The printed context or scanned image is converted into computer recognition format by using Optical Character Recognition (OCR) so that it can increase the speed of operation[3]. In model identification, all the characters are bounded and detached and then the end character image is directed to a pre-processor for removing the noise. All the characters are compared with a database of identified characters which are assembled together to form initial text pattern. The output is then given to the e-speak engine to convert text to speech and this output is given to blind users through earphones. The obstacle in the process is identified by using PIR sensor.

285 million people are evaluated to be visually impaired worldwide, 39 million are blind and 246 million have low vision[4]. This paper mainly offers a low cost system to help the blind persons. The latest progress in digital cameras, portable computers helps in designing the camera based products that combines the computer vision technology and the optical character recognition system. Software's such as video magnifiers, screen readers and optical aids are available to help the blind people and those with vision loss to use a computer. There are only few devices that can offer good access to blind users to read printed text in outside world. In today's world, number of challenges have been faced by the blind people because printed text is everywhere in the form of receipts,

*Author for correspondence

bank instructions over the medicines etc. When blind people are aided to read printed text and product labels, it will increase their independent living as well as foster social and economic self-efficiency.

Existing systems such as movable bar code readers constructed to name various products in an extended data base can facilitate the blind users to route the information about these products over speech[5,6]. On the other hand a gigantic drawback in that is it is extremely tough for blind users to locate the place of the barcode and exactly point the barcode reader at the barcode. Another reading assistive system such as pen scanner is developed in these analogous circumstances[7]. Other major problems for blind people are obstacle detection and identifying different denominations of the currencies. There are almost 50 currencies all over the world and each of them looking totally different. In the prototype system, a camera is used for reading the text and the currency. Text is converted into different languages such as English, Tamil, Telugu, Urdu[8,9]. The entire application was run on Raspberry Pi board.

2. Text to Speech Conversion

2.1 Software and Hardware Specifications

To implement text to speech conversion, a platform has been selected which include Operating system (Linux), Language (python), Library (Open CV (Linux-library)), CPU: 700 MHZ low power, ARM1176JZ-F applications processor, 512 MB SDRAM, micro SD of 8 GB, multi-channel HD audio over HDMI, USB 2.0 and supported 640×350 to 1920×1200 including 1080 p.

2.2 Raspberry Pi

The Raspberry Pi board used for development and experimenting the methods presented in this paper is shown in the Figure 1. It is a mini computer which is in credit-card size that plugs into a TV or monitor and uses a mouse and keyboard. It allows people of all age groups to use a computer and help them in learning programming languages like python and scratch.

2.3 Block Diagram

The text to speech framework consists of three essential mechanisms Image capturing, Data managing and speech output. The image capturing element gathers the images which contains text in the form of video or image. In our prototype system images are captured by using a WEB CAM. Using the camera, image of the object from the cluttered backgrounds or other surroundings are extracted. Text localization algorithm is used for acquiring the text from images; an Ada boost learning model is applied to localize the text in the image. Text recognition is implemented to convert the image based text to readable codes. In the prototype developed a laptop is used for data processing. The audio output is delivered to the blind users through e-speak engine and the audio is presented in English language. The block diagram that shows various parts of the system for implementing the proposed methods has been shown in the Figure 2.

Figure 1. Raspberry Pi board.

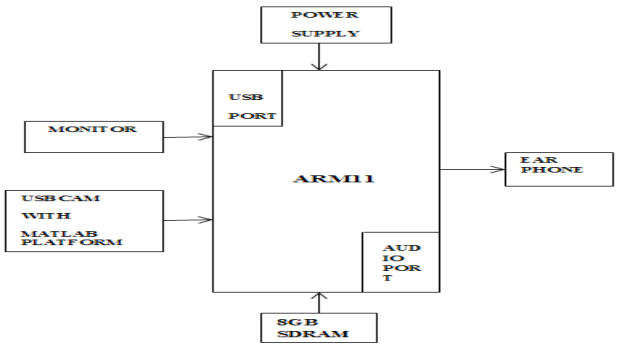

Figure 2. Proposed block diagram.

2.4 Character Recognition System

Images for Character Recognition (CR) system have been acquired by capturing the photograph of a document or product or by scanning the handwritten text or by immediately writing into a computer system. The noise available in the scanned object is removed at pre-processing stage at the time of text creation, appropriate filters such as min-max filter, Gaussian filter etc. are used to eliminate the noise. "Binarization" method changes black and white

or coloured image to binary image which is in the form of black or white. Generally grey image is the combination 0s and 1s and the binary image is the combination of 0s or 1s. The positive values above the threshold level are taken as 1s and the background with negative values are taken as 0s. Almost image processing method output is in the form of binary image. On thresholding, a grey level image with pixel values ranging between 0 and 255 represent the binary image[10]. While scanning the text, it may or may not be exactly horizontally placed, by using the slant angle correction it is exactly aligned. If the input image is too large then it is resized to diminish the dimensions to upgrade the speediness of processing. During segmentation all lines are separated by applying row histogram and by using column histogram, words are extracted from each row and last characters are extracted from words. Feature extraction is an essential part of any pattern recognition application and features of individual characters are extracted, wavelet based multiresolution technique for feature extraction has been used. Classifiers match the input characteristics with gathered sample and identify perfect equivalent input. Post processing improves the accuracy of recognition and therefore needs to be used sometimes. The step by step process used for recognizing the characters has been shown in the Figure 3.

Image Problem

Figure 3. Character recognition systems.

3. Object Detection

In prototype system presented, object is detected by using PIR sensor; it is connected to the Raspberry Pi board. When any body's hand is placed in front of that PIR sensor, e-speak engine alerts the blind person that a person is ahead[11,12].

3.1 Passive Infrared (PIR) Sensor

It is a pyro electric machine that identifies movement by computing variations in the infrared intensities by nearby objects. This motion can be sensed by examining the strong signal on a single input/output pin. It has the ability to reliably separate movable bodies from other objects as well as from stationary bodies. It is very small in size that makes it make it easy to conceal in security applications also. It has 3.3 and 5 v operation with <100 μA current consumption.

4. Currency Detection

The system presented in this paper is mainly based on image processing, filtering, edge detection, segmentation. To make the system more comprehensive, it is necessary to create a database for storing the characteristics of currency. In the system presented, Indian currency is used as example. The system is programmed using MATLAB. Several steps of processing have been implemented which include: 1. Reading the image obtained from the scanner and it is in JPEG (Joint Photographic Experts Group) format. 2. Pre-processing the image, removing the noise included and smoothening. 3. Edge detection, segmentation, pattern matching. Pre-processing of the image is done at low level of abstraction, so histogram equalization is used for increasing the image clarity by modifying the brightness and contrast. By comparing the currency with database, fake currency is also determined by using the PCA algorithm. This algorithm mainly used to extract the non-linear characteristic between different variables and separate the key features of data[13].

The hardware connections that have been established using the experimental model for achieving the currency detection and for converting text to speech is shown in the Figure 4.

The hardware design includes Memory card slot of 8 GB, Power supply, Mouse and keyboard connection, Ear phone, HDMI VGA cable connected to monitor and

Raspberry pi board, USB-2 serial port, MAX 3232 and Web CAM. The Input Image, the image taken from camera and the character recognition obtained as output from the Experimental model is shown in Figure 5, Figure 6 and Figure 7 respectively.

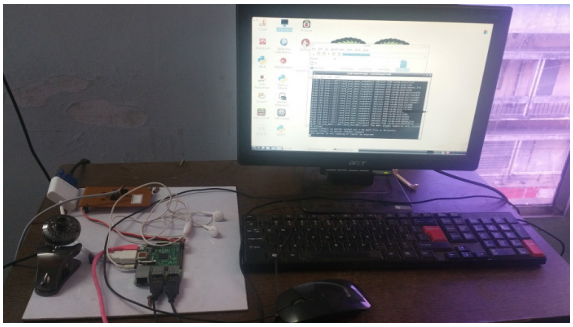

Figure 4. Hardware connections for text to speech.

Figure 5. Input images.

Figure 6. Image taken from camera.

The entire project runs on Linux library and MATLAB. In the above hardware connection Vcc is connected to pin 2 and ground to pin 6, transmitter connected to pin 10 and receiver to pin 8 on Raspberry Pi.

The input image taken is in grey form and then converted to binary so that character separation is done and each and every character is compared with the data contained in the database to find the authentication of the data. The data representing the image is fed to an e-speak engine to convert text to speech which is delivered to the blind user through ear phones. The hardware connections made for detecting the obstacles are shown in the Figure 8.

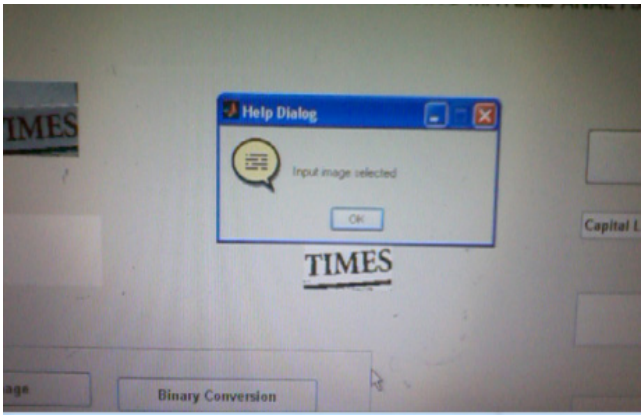

Figure 7. Character recognition.

Figure 8. Hardware connections for obstacle detection.

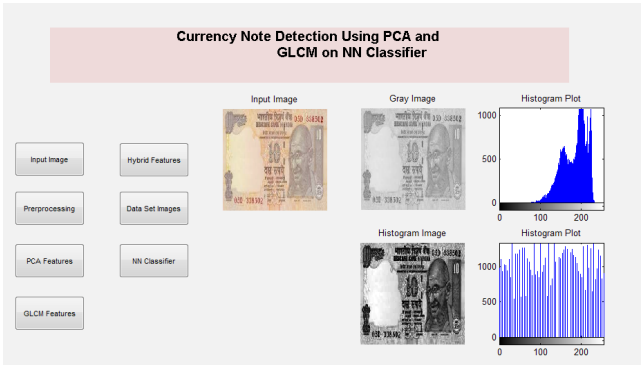

Figure 9. MATLAB executions for currency detection.

For obstacle detection PIR sensor is connected to the Raspberry Pi board when the obstacle is placed in front of

this sensor it gives information to the blind person that "A Person Is Ahead". The range of this PIR sensor is 7 meters. The hardware connection for currency and text to speech is same. The input image for currency detection is shown in Figure 9.

5. Conclusions

The proposed system helps the vision impaired people in their daily activities. This model helps to read the text and converts it in to audio form. It also helps to find out the obstacle by using PIR sensor and to find out the currency through use of a database. The PIR sensor alerts the blind user about the obstacle but cannot name the object. In future it can be further extended by using the latest obstacle detecting technologies and also helps in identifying the currency in real time.

6. References

1. Yi C, Tian Y, Arditi A. Portable camera based assistive text and product label reading from hand-held objects for blind persons. IEEE/ASME Transaction on Mechatronics. 2014 Jun; 19(3):808–17.
2. Raj Kumar N, Anand MG, Barathiraja N. Portable camera based product label reading for blind people. IJETT. 2014 Apr; 10(11):521–4.
3. Mohammad F, Anarase J, Shingote M, Ghanwat P. Optical character recognition implementation using pattern matching. International Journal of Computer Science and Information Technologies. 2014; 5(2):2088–90.
4. World Health Organization. 10 facts about blindness and visual impairment. 2015. Available from: http://www.who.int/features/factfiles/blindness/blindness_facts/en/
5. Reading ID barcodes with mobile phones using deformable templates. 2015. Available from: http://a4academics.com/be-seminar-topics/17-be-it-cse-computer-science-seminar-topics/426-reading-1d-barcodes-with-mobile-phones-using-deformable-templates
6. Introduction to Braille language. 2016. Available from: http://www.acharya.gen.in:8080/disabilities/br_intro.php
7. Pen with scanner. 2016. Availale from: http://www.prv.se/en/patents/why-apply-for-a-patent/examples-of-patents/pen-with-scanner
8. Rajendran V, Bharadwaja Kumar G. Text for developing unrestricted tamil text to speech synthesis system. Indian Journal of Science and Technology. 2015 Nov; 8(29):1–10.
9. Jyothi J, Manjusha K, Anandkumar M, Soman KP. Innovative feature set for machine learning based Telugu character recognition. Indian Journal of Science and Technology. 2015 Sep; 8(24):1–7.
10. Binary Images. 2016. Available from: http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL_COPIES/OWENS/LECT2/node3.html
11. Sensing Motion with Passive Infrared Sensors. 2016. Available from: http://www.digikey.com/en/articles/techzone/2012/jun/sensing-motion-with-passive-infrared-pir-sensors
12. Raj KM, Raj MS. Person identification for visually impaired using lab VIEW. Indian Journal of Science and Technologies. 2015 Nov; 8(31):1–4.
13. Vashishtha V, Md S. A paper currency recognition system using image processing to improve the reliability with PCA method. International Journal of Engineering Sciences and Research Technologies. 2015 Jun; 4(6):172–5.