IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 5, May 2016

System for Reading the Product Labels to Help Blind Persons

Payal V. Gulhane[1], Vrushali G. Raut[2]

M.E. student, E&TC, SCOE, Pune, India[1]

Professor, E&TC, SCOE, Pune, India[2]

Abstract: The paper proposes system gives a strategy to help blind persons by reading the product labels. In this system, one camera is used for capturing the product images. Blind person will only have to show product label in front of the camera, then the camera will capture the image. Raspberry pi 2 is used as a processor to process the image. System extracts the text from image by using SIFT (scale invariant feature transform) features of the image. KNN (k nearest neighbor) algorithm is used for calculating the Euclidean distance. The extracted text present at the output of the processor.

Keywords: KNN (k nearest neighbor), raspberry pi 2, SIFT (scale invariant feature transform).

I. INTRODUCTION

Mobile devices, wireless communication and cameras are the important things in our daily life. Because of this, new technologies are rising in the field of image processing. Cameras have important information in the captured images. The text from the images can be extracted which will be helpful to visually impaired people. The extraction of the text from the captured image is the challenging part, but many algorithms are developed to solve that problem. Proposed system makes use of SIFT features of the image and KNN algorithm to read the product labels. Algorithm extracts the text from the image so that the text can be read and KNN algorithm is used to calculate the distance between characters. Hichem Sahib, Lamberto Ballan, Giuseppe Serra, Alberto Del Bimbo proposed in their paper a framework that match and recognize multiple instances of number os reference logos [1].

Liang Wu, Palaiahnakote Shivakumara, Tong Lu and Chew Lim Tan proposed a new technique for detecting and tracking video text by using spatial and temporal information. To smooth the edge components, gradient directional symmetry technique is used before text detection. Denaunay triangulation is formed foe preserving spatial information which gives text candidates [2]. Chucai Yi, Yingli Tian proposed a method of scene text recognition. Discriminative character descriptor is designed by using several state of art features [3]. Kumuda T, L Basavaraj proposed the paper which gives an efficient algorithm for detecting and localizing text from captured images. It is based on text feature extraction by using first and second order statistics [4]. Chucai Yi, Yingli Tian proposed a framework for extracting text regions form complex background. Framework contains three main steps i.e. boundary clustering, stroke segmentation and string fragment classification [5]. Chucai Yi and Yingli Tian proposed a framework to detect text strings with arbitrary orientations in complex images. It consists of two steps i.e. image partition and character candidate grouping [6].

The existing system needs laptop for processing the images and extracting text form it which is not feasible for blind persons to carry laptop with them. This paper proposes a system which is feasible and handy for blind persons because raspberry pi is used as a processor.

II. COMPONENT USED

The proposed system gives an effective solution for reading text from captured images which will help blind person in identifying the things. Components used in the proposed system are:

A. Raspberry Pi 2

It is used to read the Text file & convert this file into the speech. It takes the images form camera for processing & reading the label of the product. It converts the images of label into text format. After this it is used to read the Text file & convert this file into the speech. The features of raspberry pi 2 are as follows.

- A 900MHz quad-core ARM cortex- A7 CPU
- 1GB RAM
- 4 USB ports
- 40 GPIO pins
- Full HDMI port
- Ethernet port
- Combined 3.5mm audio jack and composite video
- Camera interface (CSI)
- Display interface (DSI)
- Micro SD card slot
- Videocore IV 3D graphics core

Because it has an ARMv7 processor, it can run the full range of ARM GNU/Linux distributions, including Snappy Ubuntu Core, as well as Microsoft Windows 10 and the interface of the PIR & the relay with the GPIO pins of the Raspberry pi.

IJARCCE
ISSN (Online) 2278-1021
ISSN (Print) 2319 5940
International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 5, May 2016

B. Webcam Technotech ZB V-90
The product images & label images take by the camera. It sends the captured images towards the Raspberry pi for processing & detecting of the product name. Specifications of camera are as follows.

Image sensor: High Quality CMOS Sensor
Image control: Color Saturation, Brightness, Sharpness and Brightness is adjustable
Image Resolution: 3280*2460(maximum)
Exposure: Auto or Manual
Angle of View: 58 degree
Image format: RGB 24
Interface: USB 2.0
Anti-Flicker: 50Hz, 60Hz or outdoor
Frame rate: 30fps
Maximum power: 1W
Focus range: 4cm to infinity
Lens: Professional camera lens with 15 megapixel.

C. LCD monitor
It is used to show the resulting images and output of overall process on the LCD screen.

III.SYSTEM ARCHITECTURE

The proposed system is using Raspberry pi 2 with webcam 2.0 for capturing product labels. Fig. 1 shows the overview of system architecture.

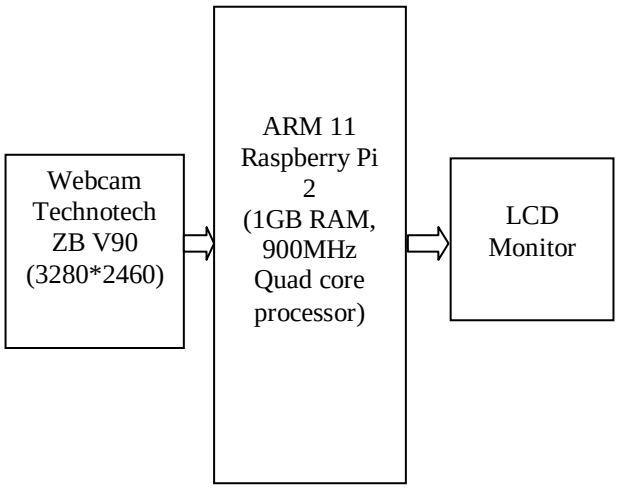

Fig. 1 Overview of System Architecture

System architecture consists of:
1. Camera / Webcam 2.0
2. Raspberry Pi
3. LCD Monitor

The product images and label images take by the camera. Camera is used for capturing the product images on which text should be read. It takes the captured images and sends it to raspberry pi. Raspberry pi 2 have ARM 11 core on it. It has processing speed of 900MHz with RAM of 1GB. In the proposed system, raspberry pi is used to read the text file and converts this file into speech. Raspberry pi 2 is the processor used for processing the image and to recognize text from image, SIFT features of the images are used. SIFT algorithm is used for logo and character recognition and KNN algorithm is used for calculating distance between them. Raspberry pi recognizes the characters from the text by using algorithm. LCD monitor is used to show the resulting images and output of overall process on the LCD screen. Fig. 2 shows the flowchart of overview of algorithm that is to be followed.

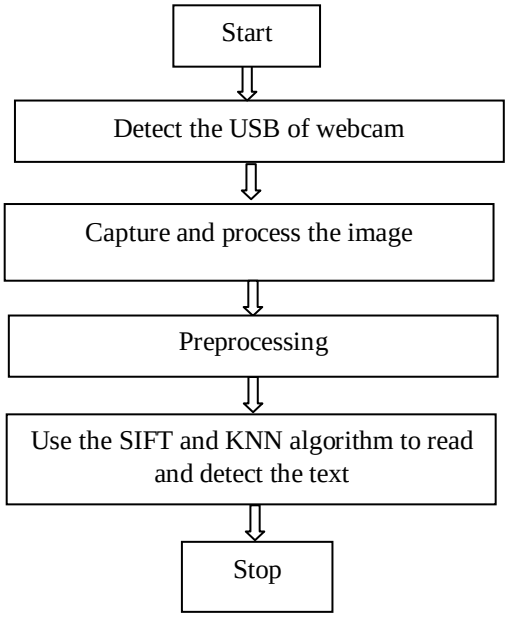

Fig. 2 Flowchart of Overview of Algorithm

Flowchart shows the step by step process performing one after the other.

IV.RESULTS

The implementation of the proposed system is difficult one. To minimize the difficulty, reading simple text from the product images is started first. Setup of overall project showing the raspberry window is as shown in Fig. 3.

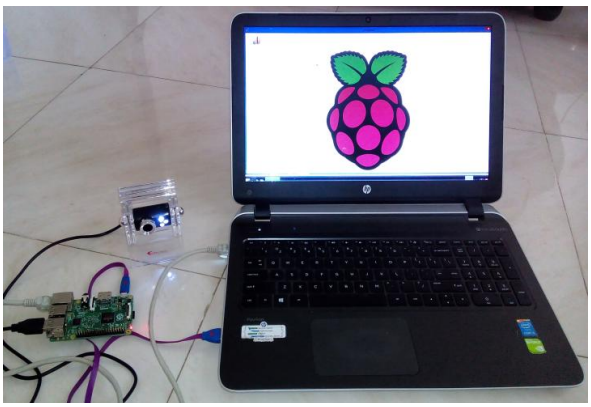

Fig. 3 Setup of Overall Project Showing Raspberry Window

For capturing the images of the product, product is placed or hold in front of the camera. The output of the captured images is shown in Fig. 4.

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 5, May 2016

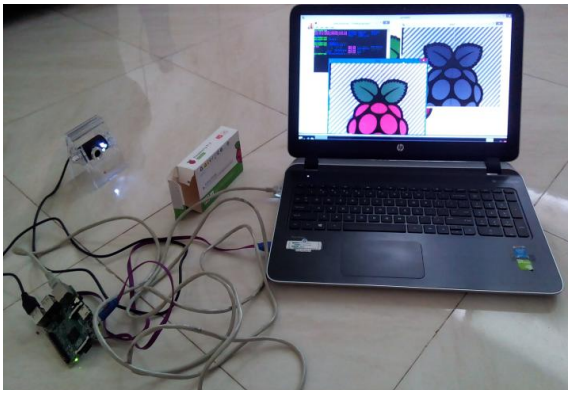

Fig. 4 Setup of Capturing an Image and Giving it in the form of RGB and Gray Format

The captured images are given in both RGB and gray format. To reduce the size and increase the speed, processes are done on gray images. One product example is taken to read the label. The result of reading the product name is as shown below in Fig. 5.

Fig. 5 Reading the Text from the Image

Fig. 5 shows how the product name i.e. lays is detected. In SIFT; features of the images are taken into considerations. SIFT works on the corner points of the images / text. Fig. 6 shows how the features of the product are matched with the database. Euclidean distance is calculated by using KNN (k nearest neighbor) algorithm.

Fig. 6 Matching the Keypoints of the Product

Label of the product is compared with the images present in the database and according to that output is given that what is the product name.

V. ADVANTAGES

Advantages of the proposed system are:

1. Provides solution to the problem faced by blind people.
2. Helps blind person to do their activities independently.
3. System is handy.

VI. CONCLUSION

Product label reading helps blind person to do their daily activities by their own. Proposed system has analyzed in all aspects regarding product label reading. It has been found that the main conceptual working of text reading is completely dependent on the algorithm that is to be used and hardware implementation of the system. Raspberry pi 2 is used to increase the speed and reduce the time. The drawback of the existing system is tried to overcome in the proposed system.

ACKNOWLEDGMENT

I would like to thank the anonymous referees for their guidance and insightful suggestions that has improved the quality of the paper. Also, I would like to thank my project guide, **Prof. V. G. Raut** for her valuable guidance.

REFERENCES

[1] Hichem Sahib, Lamberto Ballan, Giuseppe Serra, Alberto Del Bimbo, “Context dependant logo matching and recognition,” IEEE Trans. on Image Processing, vol. 22, no. 3, March 2013.

[2] Liang Wu, Palaiahnakote Shivakumara, Tong Lu and Chew Lim Tan, “A new technique foe multi-oriented scene text line detection and tracking in video,” IEEE Trans. on Multimedia, vol. 17, no. 8, August 2015.

[3] Chucai Yi, Yingli Tian, “Scene text recognition in mobile applications by character descriptor and structure configuration,” IEEE Trans. on Image Processing, vol. 23, no. 7, July 2014.

[4] Kumuda T, L Basavaraj, “Detection and localization of text from natural scene images using texture features,” IEEE int. conf. on Computational Intelligence and Computing Research, 2015.

[5] Chucai Yi, Yingli Tian, “Localizing text in scene images by boundary clustering, stroke segmentation, and string fragment classification,” IEEE Trans. on Image Processing, vol. 21, no. 9, September 2012.

[6] Chucai Yi, Yingli Tian, ”Text string detection from natural scenes by structure-based partition and grouping,” IEEE Trans. on Image Processing, vol. 20, no. 9, September 2011.