Available online at http://www.journalijdr.com

IJDR

International Journal of
DEVELOPMENT RESEARCH

ISSN: 2230-9926

International Journal of Development Research
Vol. 06, Issue, 08, pp. 9090-9093, August, 2016

Full Length Research Article

SYSTEM FOR PRODUCT LABEL READING TO ASSIST THE BLIND PERSONS

***Payal Vilasrao Gulhane and Vrushali G. Raut**

Electronics and Telecommunication, Sinhgad College of Engineering, Pune, India

ARTICLE INFO

Article History:

Received 24th May, 2016
Received in revised form
28th June, 2016
Accepted 16th July, 2016
Published online 30th August, 2016

Key Words:

Camera, Scale invarient
Feature transform (SIFT),
K nearest neighbor (KNN),
Processor.

ABSTRACT

It is very difficult for blind persons to do their daily work by their own. This paper propose a system to help blind persons read text labels and product packaging of hand held objects in their daily life. This system consists of camera for grabbing the image, processor for processing the grabbed image and headset for audio output. Blind user will only have to show objects in front of the camera and capture an image. The camera will grab the image. Input image given to the system is then converted into gray scale and apply SIFT (scale invariant feature transform) features and KNN (k nearest neighbor) algorithm on the image to detect text. Algorithm recognizes text from the image and according to that voice output of the recognized is given to the blind user. If product label is not detected, system gives instruction that product is not found.

INTRODUCTION

It is very difficult for blind person to do their work independently. To help blind people, a system is designed which help blind user to do their work by their own. The system uses camera, raspberry pi 2 and headset. The existing system uses laptop as a processor for processing the image. It is difficult for blind person to carry laptop with them all the time and it is costly also. The propose system uses raspberry pi as a processor which is a credit card sized hardware. So it is easy for blind user to carry system with them. Chucai Yi, Yingli Tian, Aries Arditi proposed in their paper motion based method to know the region of interest. Mixture of Gaussian based background subtraction method is used to extract the object (Chucai *et al*., 2014). Chucai Yi, Yingli Tian in another paper proposed framework of text string detection in two points i.e. image partition and character grouping candidate (Chucai *et al*., 2011). Sunil Kumar, Rajat Gupta, Nitin Khanna, Santanu Chaudhury, Shiv Datta Joshi proposed a technique for extraction of text areas by using globally matched wavelet filters (Sunil Kumar *et al*., 2007).

****Corresponding author: Payal Vilasrao Gulhane***
Electronics and Telecommunication, Sinhgad College of Engineering, Pune, India

Hichen Sahbi, Lamberto Ballan, Giuseppe Serra, Alberto Del Bimbo proposed a system to match reference images and logos in the database (Hichen Sahbi *et al*., 2013) S. Shamini, Dr. N. Jaisankar proposed modified context dependent algorithm which gives accuracy in logo matching (Shamini, 2014) Teofilo E. de Campos, Bodla Rakesh Babu proposed a system based on nearest neighbor and SVM classification, (Teofilo *et al.,*2009). The proposed system uses character recognition using scale invariant feature transform (SIFT) features and k nearest neighbor (KNN) algorithm for extracting text from the images. Texts present on the images are taken into consideration for output result.

MATERIALS AND METHODS

The proposed system consists of camera, raspberry pi 2 and headset. Camera is used for capturing the images, raspberry pi 2 is used for processing the captured images and headset is used for voice output. Raspberry pi 2 is used for real time operations and fast processing as it contains 1GB RAM. System architecture contains:

- Camera
- Raspberry pi 2
- Headset

This system uses a solution for detection and recognition which is context dependent similarity (CDS). CDS gives results on the basis of spatial context of local features. Context is defined in terms of single SIFT keypoint. If a set of SIFT interest point is given which is X then the context is given by set of points spatially close to X with some constrains. By using CDS, local regions are compared between images. CDS is used for finding interest points between 2 frames. For applying CDS, it requires developing matching criterion and it should verify probability of success for image recognition. For recognition of test images, some threshold point is fixed. The images are stored in the database. The image captured by the camera is taken for process. If the SIFT features are above threshold level then the text is detected. In SIFT, corner point are taken into consideration for reading the images. Algorithm of CDS recognition is as follows. Input is reference image is I_x and test image is I_y. The overall architecture of system is as shown in Figure 1.

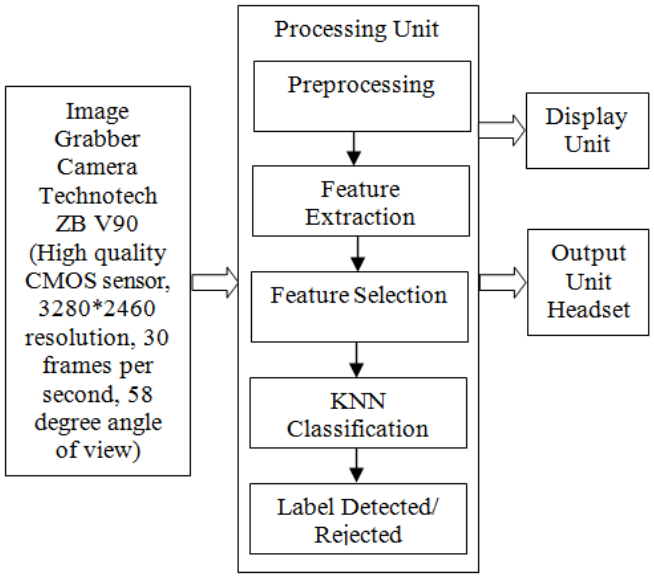

Figure 1. Overall System Architecture

K nearest neighbor (KNN) algorithm is used to calculate the Euclidean distance between characters. KNN algorithm checks the nearest neighbor in the images and compares them. Euclidean distance between points x and u is given by:

$$d(x,u) = \sqrt{\sum_{i=1}^{n}(x_i \quad u_i)^2} \quad \text{.....(1)}$$

The proposed system uses components like raspberry pi 2, camera and headset.

Raspberry pi 2

Raspberry pi 2 is used as a processor. It processes the captured image and gives in the form of text using algorithms. The features of raspberry pi 2 are as given below.

- A 900MHz quad-core ARM cortex
- 1GB RAM
- 4 USB ports
- Ethernet port
- Combined 3.5mm audio jack
- Camera interface (CSI)
- Micro SD card slot

Webcam Technotech ZB V90

Camera is used to capture images. The captured images are given to the raspberry pi 2 according to that product labels are going to read. Webcam technotech ZB V90 has the following features.

Image sensor: High Quality CMOS Sensor
Image control: Color Saturation, Brightness, Sharpness and Brightness is adjustable
Image Resolution: 3280*2460(maximum)
Angle of View: 58 degree
Image format: RGB 24
Interface: USB 2.0
Frame rate: 30fps
Maximum power: 1W
Focus range: 4cm to infinity
Lens: 15 megapixel.

Flowchart of Algorithm of the proposed system is as shown in Fig 2. The system is started with capturing the product image. After capturing the image preprocessing is done. In preprocessing image enhancement, noise removal is done. On that image feature extraction is completed by using SIFT features and then the image is filtered. KNN algorithm is applied on the filtered image to calculate the distance between points. According to the result, system gives the output that product label is rejected or detected. If the product name is read correctly then the output is given in the form voice through headset.

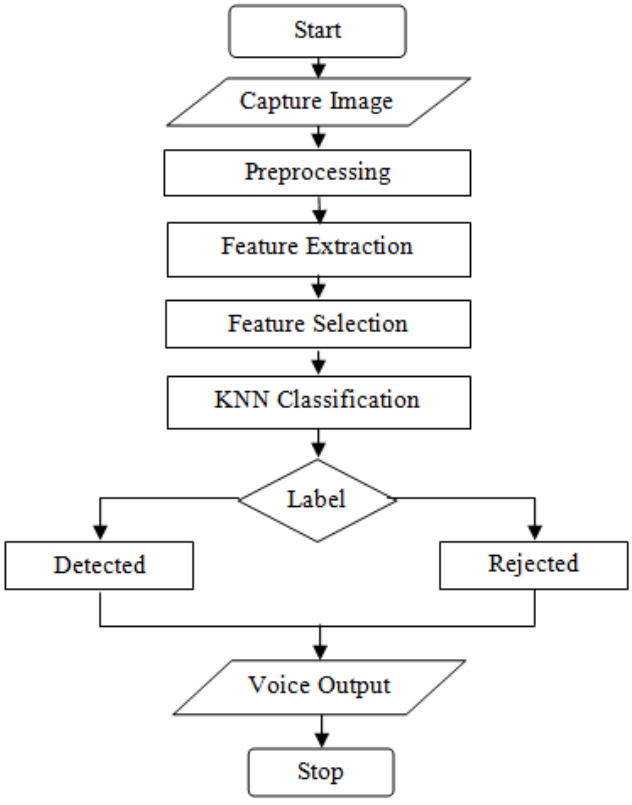

Figure 2. Flowchart of the System

RESULTS

The implementation of the proposed system is not so easy. For checking purpose one product label is taken for processing. The overall setup of proposed system contains camera, raspberry pi2 and headset as shown in figure 3. Laptop is used for showing the results.

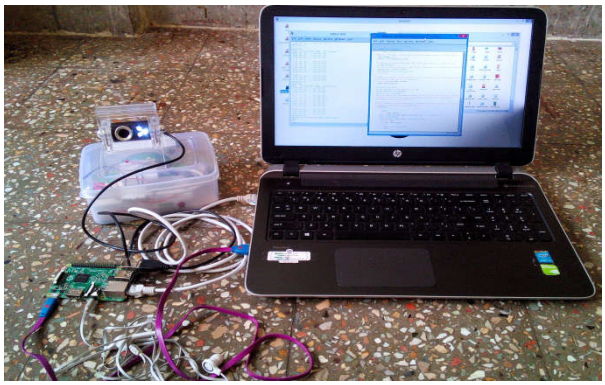

Figure 3. Overall Setup of the Proposed System

Camera is used for capturing the images of the product. Camera takes the images in two format i.e. in gray and RGB. The overall work is done on gray images because RGB images are 24 bit and gray images are 8 bit. So, it is easo to work on gray images and it also gives fast processing and requires less memory. Database is created for 10 different images and their voice in .wav format. For voice output pygame library is used. Database of the images are shown below in Figure 4.

Figure 4. Database Created for the Proposed System

One of the dettol product examples is being explained below and it shows how the working is done on the image to recognize the text. The python window and coding of the system is as shown in Figure 5. One product image is taken for checking that the system is working correctly or not. The comparison is done on the basis of two algorithms i.e. SIFT features of image and KNN algorithm. KNN algorithm finds the nearest neighbor distance and work according to the distance. Figure 6 shows the captured image of dettol product label.

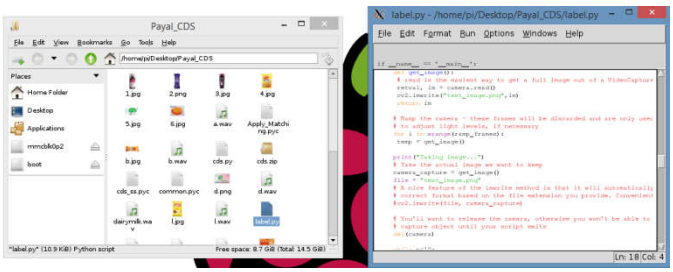

Figure 5. System Showing the Coding of Product Label Reading

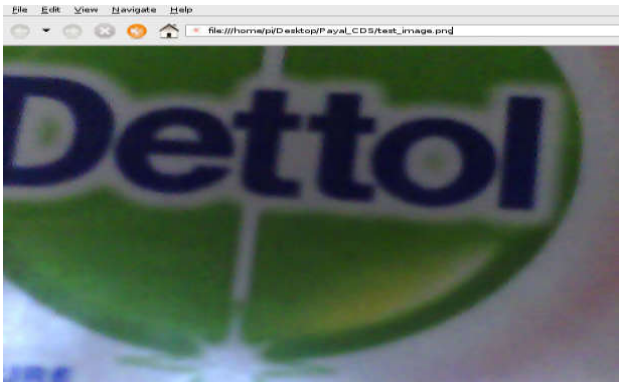

Figure 6. Showing the Captured Image of the Product

After capturing the image when the code is run, python shell window appears on the screen. It shows the comparison between image present in the database and the captured image. Figure 7 shows the result of the comparison.

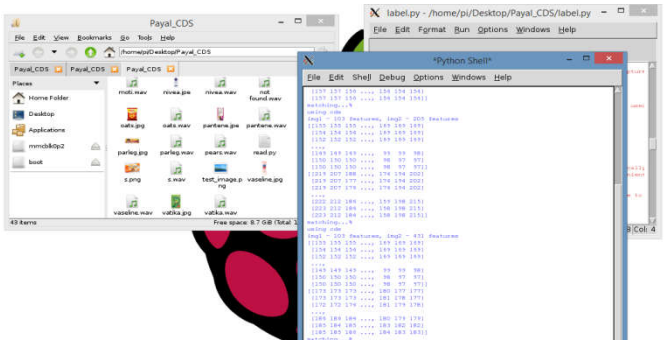

Figure 7. Window Showing the Matching of Image with the Image Present in the Database

It is compared with the SIFT features of image and KNN algorithm as measuring distance between characters is important one. The features between figure 6 are compared using SIFT and showing result in figure 7. It shows that how many features are matched and according to that voice output is given. Figure 8 shows that the product 'dettol' is read and the output is given in voice format. In this way the product label is detected and given at the output of raspberry pi 2 audio jack. The output is given in the form of voice via headset. For voice output, pygame library is used.

```
img1 - 82 features, img2 - 150 features
[[ 55  55  55 ...,  68  68  68]
 [ 55  55  55 ...,  68  68  68]
 [ 54  54  54 ...,  68  68  68]
 ...,
 [104 105 105 ...,  93  92  92]
 [104 104 105 ...,  93  93  93]
 [104 104 105 ...,  93  93  93]]
[[170 170 170 ..., 170 170 170]
 [170 170 170 ..., 170 170 170]
 [170 170 170 ..., 170 170 170]
 ...,
 [170 170 170 ..., 170 170 170]
 [170 170 170 ..., 170 170 170]
 [170 170 170 ..., 170 170 170]]
matching...%
None None 39 None None None None None None None
detol
>>>
```

Figure 8. Image Showing the Results

Conclusion

The system is used to help blind persons to read product labels means to do their work by their own. To increase the speed and storage capability raspberry pi 2 is used. As raspberry pi 2 has 1GB RAM it has maximum storage capability. Detection and recognition of images are done on the basis of SIFT features of images. SIFT features works on the corner points of the images. KNN algorithm is used to calculate the Euclidean distance. Database is created for different types of images.

REFERENCES

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

Chucai Yi, Yungli Tian, 2014. "Portable Camera Based Assistive Text and Product Label Reading From Hand Held Objects for Blind Persons," IEEE Transactions on Mechatronics, vol. 19, no. 3, June.

Chucai Yi, Yungli Tian, 2011. "Text String Detection From Natural Scenes by Structure Based Partition and Grouping,.

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

Hichen Sahbi, Lamberto Ballan, Giuseppe Serra, Alberto Del Bimbo, 2013. "Context Dependent Logo Matching and Recognition," IEEE Transactions on Image Processing, vol. 22, no. 3, March.

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

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.

Shamini, S., Dr. N. Jaisankar, 2014. "Modified Context Dependent Similarity Algorithm for Logo Matching and Recognition," Elysium Journal, vol. 1, special issue 1, September.

Sunil Kumar, Rajat Gupta, Nitin Khanna, Santanu Chaudhury, Shiv Datta Joshi, 2007. "Text Extraction and Document Image Segmentation Using Matched Wavelets and MRF Model," IEEE Transactions on Image Processing, vol. 16, no. 8, August.

Teofilo E. de Campos, Bodla Rakesh Babu, "Character Recognition in Natural Images," Computer Vision Theory and Applications, pp. 273-280, 2009

****