Wearable Assistance Device for Visually Disabled

Shivankit Bisht
Deptt. of CSE & IT, Jaypee Institute of Information Technology, Noida, India

Gautam Khanna
Deptt. of CSE & IT, Jaypee Institute of Information Technology, Noida, India

Suma Dawn
Faculty, Deptt. of CSE & IT Jaypee Institute of Information Technology, Noida, India

ABSTRACT
There are 37 million visually disable people in this world [1]. Advances in computer human interface and 3D world perception permit production of affordable and user friendly devices to assist the visually disabled [3]. This research paper presents a different approach in wearable computing using open source software and inexpensive hardware to assist the visually disabled section of population.

General Terms
Wearable Computing

Keywords
Wearable Computing, Visually Disabled, Assistive Device

1. INTRODUCTION
Assistive devices allow people with disabilities to enhance their abilities and allow for a better participation in the society. A number of wearable assistive devices have been developed for task-specific solutions for activities such as reading and travel. Given the fact that sight is restricted, they try to open new communication channels through hearing and touch.

For example, the Huffman base 4 text entry gloves [10] uses pinches between the thumb and fingers on the user's right hand for character inputs and commands.

1.1 Significance of Assistive Devices in Current Scenario
Worldwide the visually impaired population is estimated at 37 million [1]. Majority of them reside in poor countries.

For these people lack of sight is a major barrier in daily living [2].They find difficulty in reading, writing, way finding and interaction with other people and surrounding environment. Hence, cheap and efficient devices are needed to assist this section of society. Devices like Drishti, a wireless pedestrian navigation system open new avenues to assist the visually impaired. It integrates several technologies including wearable computers, voice recognition and synthesis, wireless networks, Geographic Information System (GIS) and Global positioning system (GPS).

2. METHODOLOGY
This project employs low cost hardware like webcam and headphone to implement a simple and efficient device to assist in navigation and object detection. The device takes command from the user through microphone and then the cameras detect the object that exists in its database. When detected, the device estimates the distance of the object from camera and communicates the same to the user through auditory output.

Throughout the development of this project the main objective was to create a low cost device using common items and free open source software.

Multiple algorithms and open source tools were used to implement the project.

2.1 Speed Up Robust Features (SURF)
To implement pattern matching SURF, identifies the interest points in the first set with each interest point having unique description not depending on scale and rotation. Then, given an input image determine which object it contains.

2.2 Brute Force Descriptor Matcher
For each descriptor in the first set, this matcher finds the closest descriptor in the second set by trying each one. This descriptor matcher supports masking permissible matches of descriptor sets.

2.3 KNN Match
Knn Match Finds the k best matches for each descriptor from a query set.

2.4 Microsoft Speech API
The Speech Application Programming Interface developed by Microsoft identifies the input speech and converts into text for further processing.

2.5 Ransac
It is an iterative method to estimate parameters of a mathematical model from a set of observed data which contains outliers. It is a non-deterministic algorithm in the sense that it produces a reasonable result only with a certain probability, with this probability increasing as more iteration are allowed.

3. ARCHITECTURE
The architecture of project can be illustrated in Figure 1. The input is taken through specific voice commands. On receiving the specific command the image is taken from the webcam. The image is then run through SURF which finds the contrast points in the image. Then, Brute Force Descriptor Matcher finds the closest match with the objects already in database. Knn match finds the best k matches of the image taken. Finally RANSAC iteratively processes it and confirms the object detection.

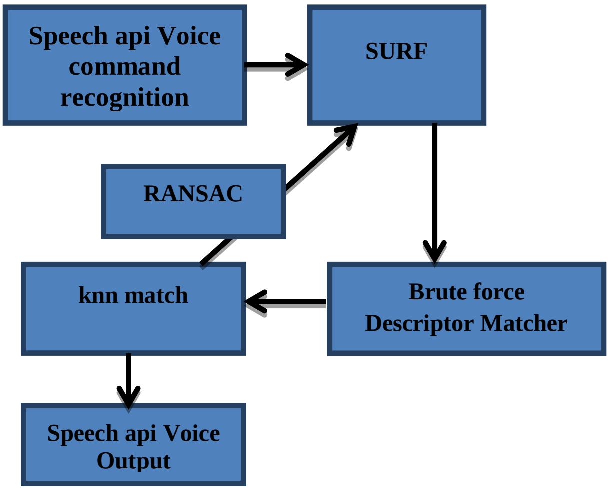

Figure 1: Project Architecture

3.1 Design & Modules

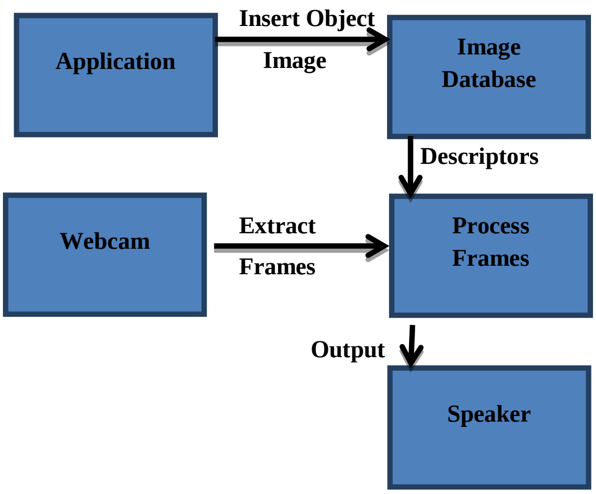

Figure 2: Project Design

3.1.1 Speech module

This is the main module of the program. Written in c#, other modules of the program are called from this module. It uses the defined speech recognition services of a windows desktop system using a "SpeechRecognitionEngine" for speech recognition and text to speech activities. It also allows the user to start the scene description and path assistance module using the specified commands in the grammar.

3.1.2 Object detection module

This module is responsible for the object detection of prototypes specified in the database. Written in Visual C++ using opencv libraries it creates a new thread for different objects in the database. Each thread then reads the images from the database, extracts the SURF key points and creates the object descriptors. These object descriptors are then matched with scene descriptors calculated from images from the webcam using a brute force matcher. If positive results are there in five consecutive frames then the object is confirmed and the user is given audio description of the object. The second webcam is used to calculate the disparity for object position to calculate an approximate distance to the object.

3.1.3 Path Assistance

This module is very similar to the object detection module except for the threshold and descriptor extraction parameters which have been optimized for the path detection markers used.

3.1.4 Distance Measurement

The distance between the object and the user is calculated using stereovision [11]. The two webcams have the same parameters, i.e. the same focal length f and the same view angle θ Image of the target T will be at distance x_1 in the left camera and at a distance x_2 in the right camera. b is the distance between the two cameras.

$$\frac{b1}{D} = \frac{x1}{f}$$

$$\frac{b2}{D} = \frac{x2}{f}$$

Since, $b = b1 + b2$

$$b = \frac{D}{f}(x2 - x1)$$

$$D = \frac{bf}{(x2 - x1)}$$

$$\tan\frac{\theta o}{2} = \frac{xo}{D} = \frac{x1}{f}$$

$$f = \frac{xo}{2\tan\left(\frac{\theta o}{2}\right)}$$

$$D = \frac{bxo}{2\tan\left(\frac{\theta o}{2}\right)(x2 - x1)}$$

Here, xo is width of image in pixels and x2-x1 is disparity between the two images in pixels. To compensate alignment errors, another term ϕ is added that is calculated experimentally.

Finally, distance is calculated using

$$D = \frac{bxo}{2\tan\left(\frac{\theta o}{2} + \phi\right)(x2 - x1)}$$

4. RESULTS

Fields of image processing and object recognition are relatively new. A lot of research is still under progress in this field. Table 1 shows the achieved results of our project. Stable object recognition and distance to object calculation can be obtained but with a huge possible margin of error because of environmental factors.

Table 1: Results

Type of Test	Total Tests	Success	Failed
Voice Recognition	40	38	2
Object Detection	40	18	22
Distance Calculation	25	23	2
Marker Detection	40	28	12

Over a set of 300 iterations, the project performance the performance of the device the device performed more appropriately.

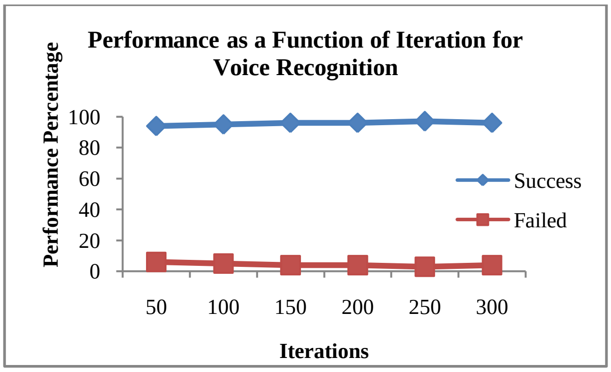

Figure 3: Performance as a Function of Iteration for Voice Recognition

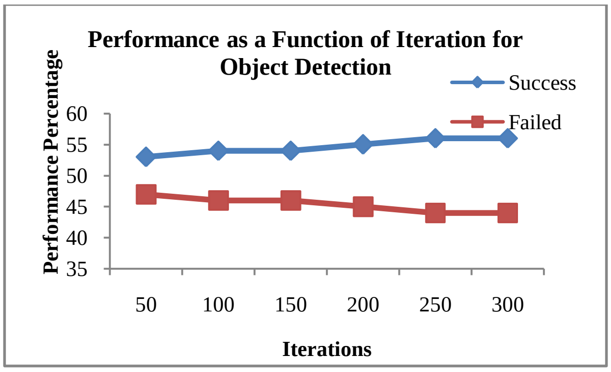

Figure 4: Performance as a Function of Iteration for Object Detection

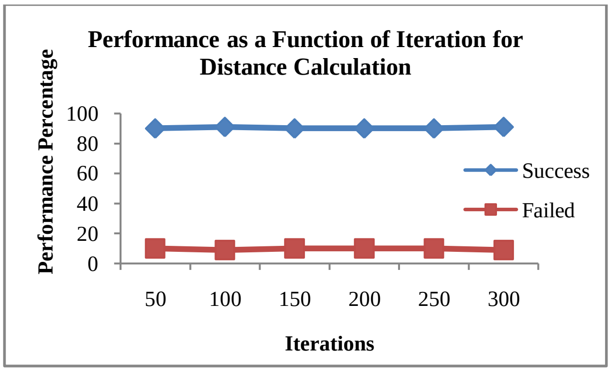

Figure 5: Performance as a Function of Iteration for Distance Calculation

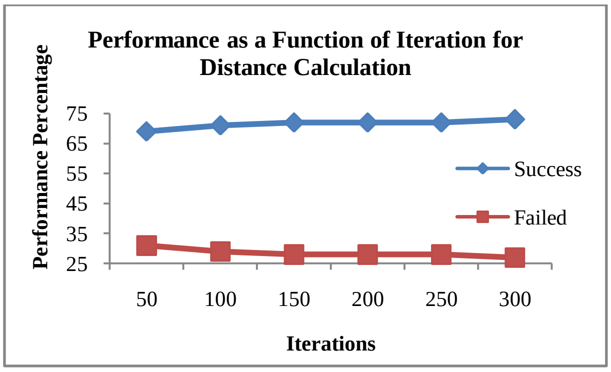

Figure 6: Performance as a Function of Iteration for Distance Calculation

The results were achieved in well-lit environment with almost zero external noise.

The false output can be attributed to following factors:

1. Environmental
2. Error in input sensors like webcam, microphone etc.
3. Calibration errors

4.1 Project Prototype and Working

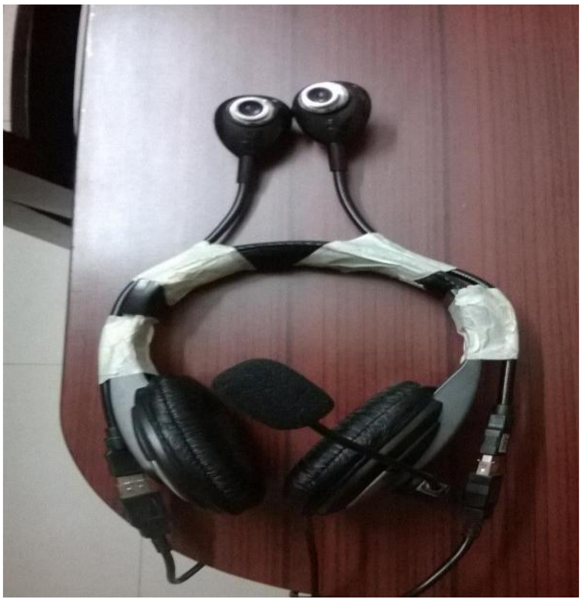

Figure 7: Project Prototype

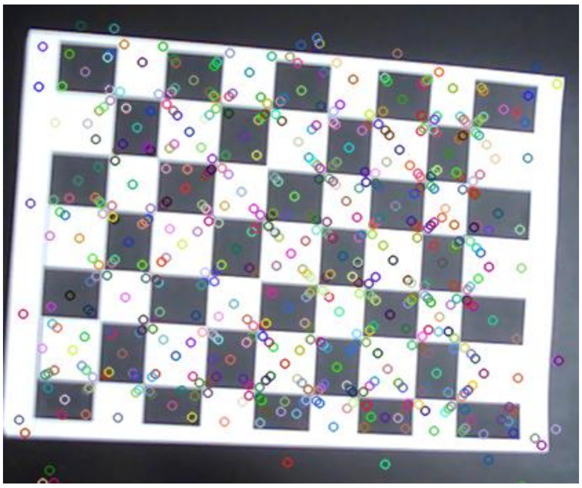

Figure 8: Object Detection

5. CONCLUSION

Object recognition and distance to object calculation using a video is very sensitive and prone to erroneous results but can be highly improved, provided proper lighting and environment conditions exist. Also, better quality of audio and video input devices may further enhance the chances of success rate. Another aspect which is presently a deterring factor for our project is an increased database set of identified objects. Though, in turn, this may increase the time requirement, the project output would also surge.

6. ACKNOWLEDGEMENT

Shivankit Bisht and Gautam Khanna would like to thank Tanu Mohan Sharma for her valuable comments and for sharing her knowledge during the development of this paper.

7. REFERENCES

[1] World Health Organization (2009) Visual impairment and blindness - Fact Sheet N°282. Available online at: http://www.who.int/mediacentre/factsheets/fs282/en/

[2] Blind World Magazine (2006) Breaking the chains of paternalism. Available online at: http://home.earthlink.net/~blindworld/NEWS/6-06-14-02.htm

[3] Brabyn J, Seelman K and Panchang S, Aids for people who are blind or visually impaired. In: Cooper R, Ohnabe H, and Hobson D (Eds.), An introduction to rehabilitation engineering, Taylor & Francis, pp 287-313.

[4] C Kirner1, T G Kirner2, R S Wataya3, 2010. Using augmented reality to support the understanding of three-dimensional concepts by blind. Proc. 8th Intl Conf. Disability, Virtual Reality & Associated Technologies

[5] Paul Viola, Michael J. Jones, 2004. Robust Real-Time Face Detection". International Journal of Computer Vision 57(2), 137–154

[6] Abdel Salam (Sumi) Helal, Steven Edwin Moore, Balaji Ramachandran, 2004.An Integrated Navigation System for Visually Impaired and Disabled

[7] David G. Lowe, 1999. Object Recognition from Local Scale-Invariant Features. Proc. of the International Conference on Computer Vision, Corfu

[8] Paul Viola, Michael Jones 2001.Rapid Object Detection using a Boosted Cascade of Simple Features. Accepted Conference on Computer Vision and Pattern Recognition

[9] Manaf A. Mahammed1, Amera I. Melhum2, Faris A. Kochery, 2004. Object Distance Measurement by Stereo Vision. International Journal of Science and Applied Information Technology (IJSAIT), Vol.2, No.2, Pages: 05-08

[10] Bajer, MacKenzie, I.S. Baljko M. 2012..Huffman base-4 Text Entry Glove (H4-TEG). 16th International Symposium on Wearable Computers

[11] Manaf A. Mahammed, Amera I. Melhum, Faris A. Kochery 2013. Object Distance Measurement using StereVISION. International journal of Science and Applied Information Technology, Vol.2, No.2,pp 05-08