Object Detection in Android Smartphones for Impaired

Prof. S.P. Jadhav[1], Sonali S. Jayswal[2], Stelly. Tommy[3], Harshvardhan D. Dhaware[4], Akshay R. Vijapure[5]

Professor, Computer Engineering, Department, PVPIT, Pune, India[1]

Student, Computer Engineering, Department, PVPIT, Pune, India[2,3,4,5]

Abstract: The Main goal is to eliminate the need of dedicated devices and other wearable devices. We are implementing to help the visually impaired people using simple android Application with the help of smart phones. By using image processing, object detection, speech synthesis the highlighted key components to use this application. Detecting object using image processing technique can be used in multiple industrial as well as asocial applications. The project purpose is to use object detection for blind people and give them audio or vocal information about it as well as we are doing feature extraction to search for objects in a camera view. To improve the day to day lives of visually impaired blind people and people with low vision. Keeping in mind the challenges faced by them we felt the need of an Android application to provide them necessary information about the environment and the objects. The paper gives an overview of a general method of object recognition and significance of ORB over SIFTS and SURFS in different cases. This paper also provides an idea to implement ORB algorithm on FPGA to increase the execution speed by utilizing the reconfigurable nature and pipelining of the FPGA.

Keywords: Image Processing, Speech Synthesis, Object Detection, Android Application.

I. INTRODUCTION

Object detection is a largely upcoming scope for detecting object using image processing techniques which are used for social and industrial applications. Purpose of this project is object detection for visually impaired by using speech synthesis and feature extraction in camera view. Thus we conclude that object detection is useful in day to day life of visually impaired users.

1.   Image processing:-
Image processing is a conversion of image into digital format and performs operation on it to get enhanced information and extract useful information. It is usually present in form a video frame or photograph and the result is image. We treat image as 2 dimensions and apply defined processing methods to it. It follows operations like virtualization for clarity of image, image sharpening for better sharpened image.

2.   Speech synthesis:-
Speech synthesis is the artificial production of human speech. Synthesized speech can be created by concatenating pieces of recorded speech that is stored in database. This is generated after data extraction and detection in speech synthesis. The output will be in Form of speech which will give the impaired users an idea of object detected.

3.   Object detection:-
Object detection can be implemented using BLOB and ORB algorithm. Here our aim is to detect the object from the surrounding and store them in database for future use. Once the user is motion and the surrounding object is detected the particular object is matched with object stored in the database.

4.   Android application:-
Android application is software that runs on android platform. Because the android application is built for mobile devices a typical android application is designed for smartphones or tablet pc that runs on android OS. Android application is very easy, convenient, portable and impressive in use. Also because of its feature and operation abilities more functions could be performed. It is a very powerful and upcoming platform.

II. LITERATURE SURVEY

1.   Title: Object recognition with ORB and its Implementation on FPGA Author:A.V.Kulkarni, J.S.Jagtap,Issue-11 September-2013 Desc:This paper gives an overview of a general methods of object recognition and significance of ORB over SIFT and SURF in different cases. This paper also provides an idea to implement ORB algorithm on FPGA to increase the execution speed by utilizing the reconfigurable nature and pipelining of the FPGA.

2.   Title: Object recognition in a mobile phone application for visually impaired users Author:Jae Sung Cha, Dong Kyun Lim and Yong-Nyuo Shin June, 2013 Desc:In this paper they have described main features of software

modules developed for Android smartphones that are dedicated for the blind users. The main module can recognise and match scanned objects to a database of objects, e.g. food or medicine containers. The two other modules are capable of detecting major colours and locate direction of the maximum brightness regions in the captured scenes. We conclude the paper with a shortsummaryofthetestsofthesoftwareaidingactivitiesofdailylivingof a blind user.

III. OBJECTIVE AND STUDY

This application will reduce the use of multiple devices for object detection and object recognition. This could be accomplished by providing blind people with commands about the things and movements which should be taken all over surroundings.

IV. DATA AND METHODOLOGY

DATA:

In our project we are using single core processor which will fulfil all the requirements needed for the android application. Blind user generate the video which will be divided into frames then image enhancement that is edge detection will be done after that key matching will be processed with the database then text to speech conversion using speech synthesizer which gives the direction of command. The general data structures are:

1. Internal software data structure
2. Global data structure
3. Temporary data structure
4. Database description

Internal software data structure:-
User will login and the video will be captured and templates are generated using image enhancement and edge detection process is proceeding.

Global data structure:-
Sessions: In this case the key matching will be according to key matching will be done with predefined database and text to speech conversion using voice command

Temporal data structure:-
Object is detected in recognition phase.

Database description:-
Inthedatabasewhichweareusingallthedatarelatedtoobjectswithitsallviews like front, back view are stored which is use for key matching.

METHODOLOGY:
The image which is captured is converted into grey scale which results into image enhancement using BLOB i.e. the algorithm used for object detection, ORB algorithm, and speech synthesizer provides text to speech recognition conversion which will give overall completion for application execution. Here we are showing six methodology flows.
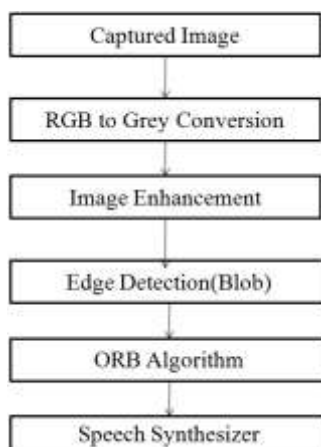

Fig: methodology

1. After starting of application video will get generate and it will get divided into frames.
2. In second phase RGB colour image will get converted into grey conversion i.e. values between 0 and 1.
3. InthirdphaseImageenhancementwhichwillhighlightalltheexistingpixelsfrom edge.
4. In fourth phase Edge detection is thoroughly dependent on above phase.
5. In fifth phase ORB algorithm uses key matching is used.
6. In sixth phase conversion of text to speech takes place by speech synthesizer.

V. MODELLING

ORB:-

An openCV enthusiast, the most important thing about the ORB is that it came from "OpenCV Labs". This algorithm was brought up by Ethan Rublee, Vincent Rabaud, KurtKonolige and GrayR. Bradski in this paper **ORB: An efficient alternative to SIFT and SURF** in 2011. As this title say it is a good alternative to SIFT and Surf in computation cost, matching performance and mainly the patents.

ORB is basically a fusion of FAST keypoint detector and BRIEF descriptor with many modifications to enhance the performance. First it uses FAST to find keypoint s, and then apply Harris corner measure to find the keypoints. It computes the intensity weighted centroid of the patch with located corner at centre. The direction of vector from this corner point to centroid gives the orientation.

To improve the rotation invariance, moments are computed with x and y which should be in a circular region of radius r. where r is the size of the patch.

Now for descriptors, OB use BRIEF descriptor, but we have already seen that BREIF performs poorly with rotation. So what ORB does is to "steer' BRIEF according to the orientation of keypoints. For any feature set of n binary tests at location (x_i, y_i), defines a $2 \times n$ matrix S which contains the coordinates of this pixels. Then using the orientation of patch, θ, its rotaion matrix is found and rotates the S to get steered version.

ORB discretises the angle to increments of $2\pi/30$ (12 degrees), and constructs a lookup table of pre-computed BREIF patterns. As long as the keypoints orientation is consistent across views the correct set of points S_θ will be used to compute its descriptor.

Very less literature is present as far as ORB is concerned. This section describes the ORB and how it differs from others. ORB is a very fast binary descriptor based on BRIEF, which is rotation invariant and resistant to noise. It can be demonstrated through experiments how ORB is at two orders of magnitude faster than SIFT, while performing as well in many situations. Authors have tested the efficiency of ORB on several real-world applications, including object detection and patch-tracking on a smart phone.

oFAST: FAST Keypoint Orientation
FAST features are widely used because of their computational properties. However, FAST features do not have an Orientation component. oFAST is the efficiently computed orientation added to the FAST.

FAST Detector FAST takes one parameter, the intensity threshold between the centre pixel and those in a circular ring about the centre. FAST does not produce a measure of corniness, and it has large responses along edges. Harris corner measure to order the FAST keypoints is employed. For a target number N of keypoints, it sets the threshold low enough to get more than N keypoints, then orders them according to the Harris measure, and picks the top N points. FAST does not produce multi-scale features. A scale pyramid is employed of the image, and produces FAST features (filtered by Harris) at each level in the pyramid. ORB uses a simple but effective measure of corner orientation, the intensity centroid. The intensity centroid assumes that a corner's intensity is offset from its centre, and this vector may be used to impute an orientation. The moments of a patch can be defined as:

$$m_{pq} = \sum_{x,y} x^p y^q I(x,y)$$

And with these moments we may find the centroids:

$$C = \left(\frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}} \right)$$

We can construct a vector from the corner's centre, O, to the centroid, OC. The orientation of the patch then simply is:

$$\theta = \text{atan2}(m_{01}, m_{10})$$

Where atan2 is the quadrant-aware version of arc tanthe centroid method is compared with two gradient based measures, BIN and MAX. In both cases, X and Y gradients are calculated on a smoothed image. MAX chooses the

largest gradient in the keypoint patch; BIN forms a histogram of gradient directions at 10 degree intervals, and picks the maximum bin. BIN is similar to the SIFT algorithm, although it picks only a single orientation.

rBRIEF: Rotation-Aware Brief

The BRIEF descriptor is a bit string description of an image patch constructed from a set of binary intensity tests. Consider a smoothed image patch, p. A binary test τ is defined by:

$$\tau(p; x, y) = \begin{cases} 1 : p(x) < p(y) \\ 0 : p(x) < p(y) \end{cases}$$

Where p(x) is the intensity of p at a point x. The feature is defined as a vector of n binary tests:

$$f_n(p) = \sum_{1 \leq i \leq n} 2^{i-1} \tau(p; x_i, y_i)$$

Steered BRIEF Matching performance of BRIEF falls sharply for in-plane rotation of more than a few degrees. A more efficient method is to steer BRIEF according to the orientation of keypoints. For any feature set of n binary tests at location (xi, yi), define the $2 \times n$ matrix

$$S = \begin{pmatrix} x_1, \ldots \ldots x_n \\ y_1, \ldots \ldots y_n \end{pmatrix}$$

Using the patch orientation θ and the corresponding rotation matrix Rθ, we construct a "steered" version Sθ of S:

$$S_\theta = R_\theta S$$

Now the steered BRIEF operator becomes:

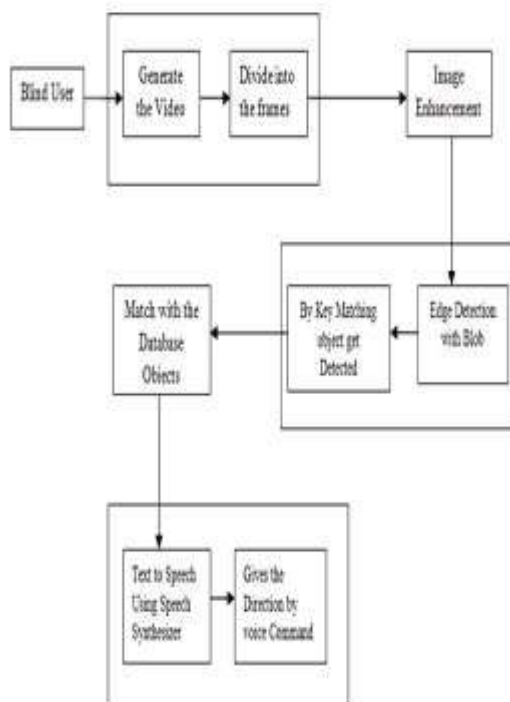$$g_n(p, \theta) = f_n(p) | (x_i, y_i) \in S_\theta$$

Application Architecture

Fig 1. System Architecture

☐This figure specifies the design that is used to solve the problem of product. We are detecting an object using the mobile camera and giving voice instructions about the direction of an object. We need to train the system first about the object information. We are then doing feature extraction to search for objects in the camera view.

☐Blind user generate the video which will be divided into frames then image enhancement that is edge detection will be done after that key matching will be processed with the database then text to speech conversion using speech synthesizer which gives the direction of command.
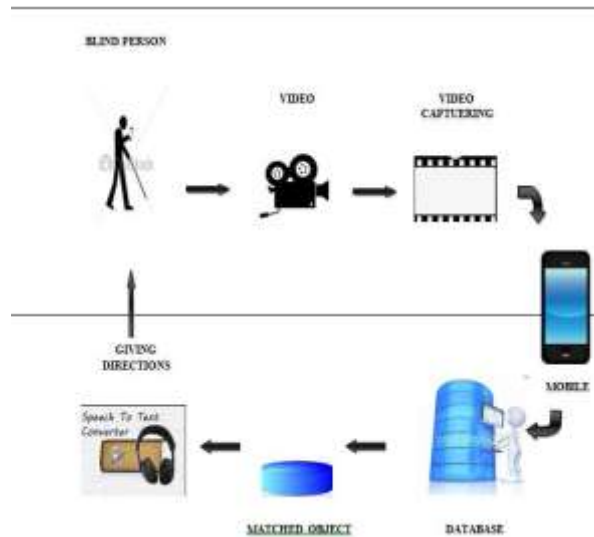
Fig:2 Client Server Architecture

Client side:

☐The user that is blind person will use smart android phone to locate the directions and the main component that is after starting the android application it will give current location of blind person and android phone camera will capture the live surroundings and it will get divided into several frames. The exact object captured which is acting as obstacle in blind person path will get send to the server side.

Server side:

☐On the other hand that is on server side the object which is sent by the client present in server will match with the object present in database and where as database is collection of already stored object. If object match successful speech to text conversion will take place.

IV. PSEUDO CODE

Step 1: Generate the video.
Step 2: Number of frames get generated from video.
Step 3: Convert Image into RGB to Gray scale.
Step 4: Check the object edges detected by the blob algorithm.
Step 5: By using ORB algorithm Setting we will set the key points for object then matches the object with the database object.
If (object=database)
then
make the text to speech conversion
else
go to step 5
Step 6: with the Speech synthesizer blind user will get the voice command
Step 7: go to step 4.
Step 8: End.

V. SIMULATION RESULTS

The simulation studies involve the working of android blind people application which contains different modules .In this current location of the blind user will be detected first then camera will be on and it will capture video which will be divided into the frames that the frame will be consider as image .From the image system is going to detect the object using blob detection algorithm. Once the objects are detected, edge detection algorithm is applied to it which will decide the object by bounding box.

Those objects are detected and key points are set by using ORB. Those selected objects are set with the key points and those key points are matched with the database images and if match found that object will be confirmed and system will give notification to the user or blind people. The notification will be given in voice by using voice synthesizer. This is how the application will help user to navigate from one location to another.
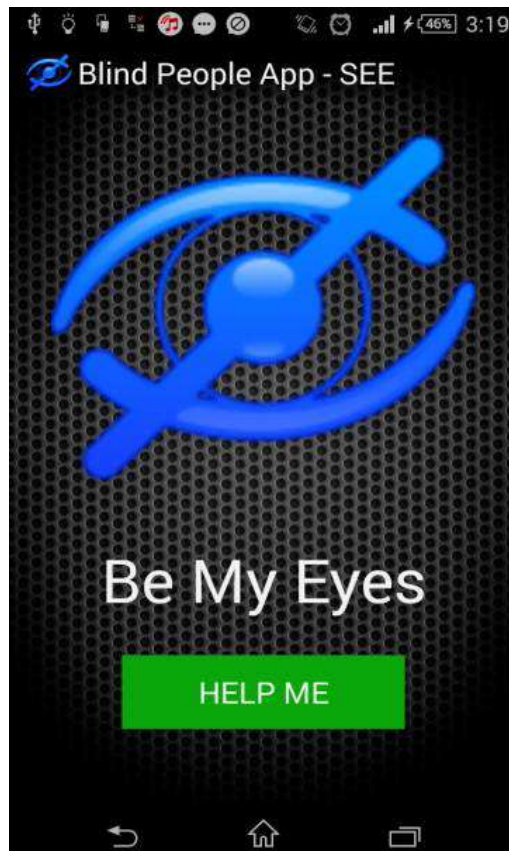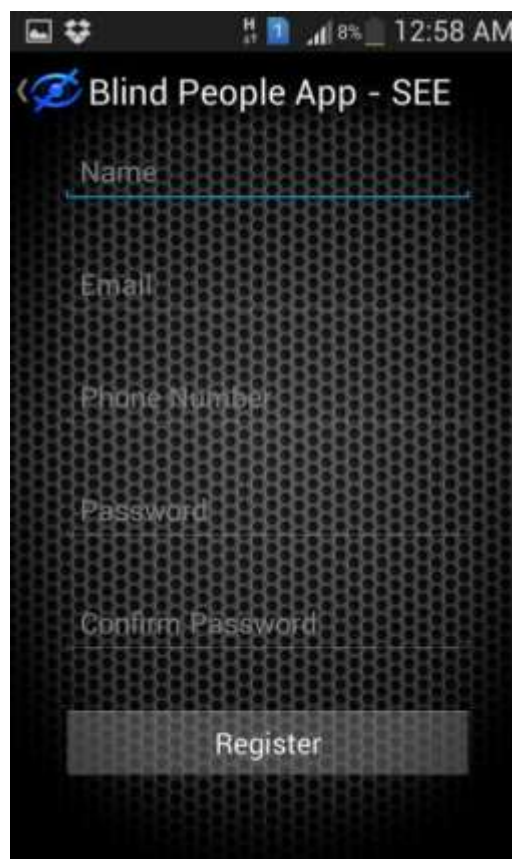
Fig 3. Application window
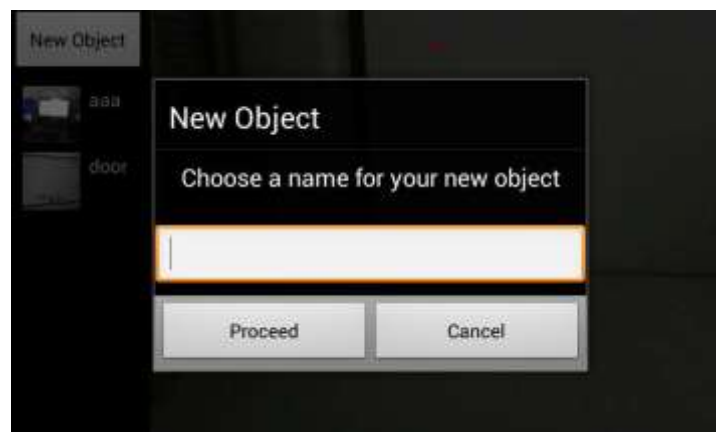

Fig 4. Login page

Fig 5. Register page

Fig 6.Creating new object

VI. CONCLUSION

This provides much important contribution. It presents a survey of many technologies that can be used by visually impaired in identification task. According to this we conclude that human powered services technology would be the winner due to its convenience and effectiveness.

ACKNOWLEDGMENT

REFERENCES

[1]   Wang Yu, Object Detection System", International Journal of Scientific and Engineering Research Volume 3, Issue 5, May-2013 1 ISSN 2229-5518.

[2]   Ananda Kanagaraj S,\ A mobile application for personal safety",9th IEEE International Conference on Collaborative Computing: Networking, Application and Work-sharing(CollaborateCom 2014).

[3]   http://en.wikipedia.org/wiki/Cheeka

[4]   K.Matusiak, P.Skulimowskiand P. Strumio, Object recognition in a mobile phone application for visuallyimpairedusers,978-1-4673-56374/13/31.002014IEEE

BIOGRAPHIES

Mr. Akshay R. Vijapure is a Student of B.E. Computer Engineering, TSSM's Padmabhooshan Vasantdada Patil Institute of Technology, Savitribai Phule Pune University, India.

Ms. Sonali S. Jayswal is a Student of B.E. Computer Engineering, TSSM's Padmabhooshan Vasantdada Patil Institute of Technology, Savitribai Phule Pune University, India.

Mr. Harshavardhan D. Dhawareis a Student of B.E. Computer Engineering, TSSM's Padmabhooshan Vasantdada Patil Institute of Technology, Savitribai Phule Pune University, India.

Ms. Stelly Tomy is a Student of B.E. Computer Engineering, TSSM's Padmabhooshan Vasantdada Patil Institute of Technology, Savitribai Phule Pune University, India.