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

*International Journal of Advanced Research in Computer and Communication Engineering*
*Vol. 4, Issue 4, April 2015*

# Object and Scene Recognition in an Android based Mobile Application

**Ms. Sonali Bhandari[1], Ms. Sonali Maid[2], Ms. Sonam Malve[3], Mr.Hardik Thakkar[4], Mr.Nilesh Sonawane[5]**

Student, IT Department, PVG's COET , Pune, India[1,2,3,4]

Professor, IT Department, PVG's COET , Pune, India[5]

**Abstract**: Around 285 million people are estimated to be visually impaired worldwide out of which 39 million are blind and 246 have low vision. About 90% of the world's visually impaired live in low-income setting and 82% of people living with blindness are aged 50 and above. So this project aims to improve the day to day lives of the blind user by providing them some information about the environment. Keeping in mind the challenges faced by the blind we had come up with an android application [1]. In this paper we present the main features and appearance of software modules dedicated to the aid of visually impaired or blind users. The main aim of developing this Application is to reduce and collapse the need of separated devices for object recognition and detection. The software modules are designed for Android operating system, used in majority of the smart phones today. There are two main trainable and educable modules, namely, the object recognition module and the Object detection module. Image Processing (Euclidean) algorithms used to identify the objects. Notification to the users is given by means of voice actions.

**Keywords**: Image Processing, Motion Detection, Object Recognition, Visually Impaired.

## I. INTRODUCTION

The visually impaired and blind people face abundant dispute in routine jobs such as making coffee or crossing roads. Identification of objects and motions in the surroundings is a primary challenge for them which normal seen people take for granted. We aim to provide some help to them via software based on android platform which will help lack of difficulty for performing the task of object Recognition and detection. Currently, the most popular choice of smart phones among visually impaired users is iPhone or Android based phones. Usually, the non operating system devices are not approved by blind users as they do not offer special functions such as text to speech conversion.

This invention specially relates to the development of technology to comfort the blind and impaired person. Our Application is one kind of the alternative for their eyes. The blind and visually impaired people face various kinds of problems in their daily life while travelling in home or at public places. However, blind and visually impaired peoples do not enjoy as many facilities as their seen counterparts. So in order to give social and normal life to blind people ,we report a solution which can be helpful to them while detecting objects like pen, pencils etc. and scenes like classrooms, labs etc. Android is popular with technology companies which require a ready-made, low cost and customizable operating system for high-tech devices. This system makes use of image recognition principle running on android system Smartphone and appropriate response is given to the user by verbal speech. The application will work on android operating system Honeycomb (3.0) onwards till Kit Kat (4.4). The scope of this project is implementation of the system for detecting different objects as well as scenes using image processing techniques. Some of the highlighted use cases can be recognition of scenes and objects for blind user. Detection of object is done in two phases, training phase and recognition phase. In training phase important part is tagging in which tags are really nothing more than keywords used to describe a piece of data. These tagged images are stored in the database for further recognition process. Once objects detected are saved and available for the further use.

the conference website. For questions on paper guidelines, please contact the conference publications committee as indicated on the conference website. Information about final paper submission is available from the conference website.

### A. *Existing System*

Assigning the Smartphone market for the visually impaired is a successful task. Over the years, there have been plenty of advancement in this area with the introduction of sample applications, voice technology and on screen accessibility features but now a new aspirant into the market is seeking to help family and brood as well. Screen reader, a not-for-profit company, has now introduced the Android based Georgie app in the U.K where there 1 in every 30 people suffer from some sort of visual impairment. It's the first Smartphone that's been designed for blind people by blind people. Georgie uses fingers and gestures to let users navigate the software, just like you do normally. However, with Georgie, you hold your finger down, wait for a beep and hear the function you're touching read out loud, whether it's to make a call, send a text, or determine location. Georgie makes extensive use of voice control software [2].

*Merits Of Existing System*

Out of the box Georgie comes with features to let users

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

*International Journal of Advanced Research in Computer and Communication Engineering*
*Vol. 4, Issue 4, April 2015*

dial a number with the voice assisted touch screen, manage contacts, use speech input to send text messages and tag previous routes or hazards (like potholes or low hanging branches) using the navigation apps. A variety of additional apps are also available for purchase and bundled into three different packages, Travel, Lifestyle or Communicate, to add more functions to support different aspects of daily life that blind people may currently find challenging[1]. These bundles are available for £24.99 each and include the following extra features:

Travel

- Near me – cover pin to point places within a vicinity of few meters.
- Buses – it informs about the bus stops arriving and the upcoming ones
- Weather – its gives exact details about the weathers in the current catchment.
- Audio Player – it helps listen to music new paper and read, listen audio books.
- Voice Assistant - it works as a service assistant and answers all questioning audio format.
- Color – helps to identify color of an object
- Communicate
- Assistance – it tells the exact location.
- Audio Tagging – add a sound clip to images so they can easily be found later
- OCR – turn a photo into a text document that can be listened to and saved
- Blogs – record and broadcast audio blogs

*Demerit Of Existing System*

The existing systems have a number of flaws, which we need to be rectified so as to build a comprehensive application. Some of the disadvantages of existing systems are as follows: When someone buys their first Smartphone, the first question is usually, what apps should I get? Apps are part of what make our phones so valuable to us, from getting directions to finding a place nearby to eat, to keeping track of various items when we travel. To overcome these Problems we had developed an app to Detect and Recognize the object.

## II. AN EASY WAY TO COMPLYMOBILE ACCESS TECHNOLOGY

For most of the past few decades, mobile access technology for blind people came in the form of specialized hardware devices. Products like talking barcode readers, color identifiers, And talking Optical Character Recognition took the form of expensive, dedicated hardware costing hundreds of dollars [3]. Such devices were limited by the capabilities of automatic Technology and had limited uptake. In the past few years, many standard mobile devices have started to include screen reading software that allows blind people to use them. For instance, Google's Android platform and the Apple iPhone (starting with the 3GS) now include free screen readers [4]. Touch screen devices like the iPhone were once assumed to be inaccessible to blind users, but well-designed, multitouch interfaces leverage the spatial layout of the screen and can even be preferred by blind people [5]. With the availability of an accessible platform, a number of Applications were developed for blind people, including GPS Navigation applications, OCR readers, and color recognizers. One of the most popular is LookTel1, which simply identifies U.S. currency denominations (but does so very reliably).

### *A. Detection Of Object And Scene Captured By Camera*

The aim was to design an application which would allow detecting and recognizing objects and scenes from images captured by the camera of a mobile device. Object and scene detection done in two phases first is image Training phase and second is Recognition phase. In training phase normal person needs to train all objects and scenes. Image tagging is important step in detection of object. Tagging refers to attaching identified information to the object. Then the entire trained object is stored in the database for recognition purpose.

When blind user captures the image of objects then this input images are send for further processing. Initially this input image is converted into grayscale format and grayscale blurring is done. After that thresholding is done and particular threshold value is decided. In order to know the boundaries of image edge detection is performed on that image. Then this RGB images are converted into HSV format [6].Input image have several invariant features this are invariant according to illumination, viewpoint etc. This invariant features are extracted from input image and feature vectors are generated .This feature vectors are used for further recognition purpose by using serialization process and are compared to vectors from database and appropriate voice response is given to the blind user.

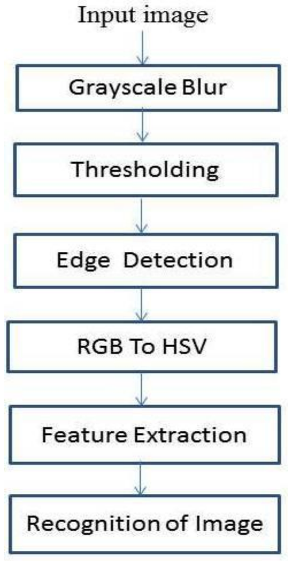


Fig.1 Flow diagram

After capturing the image of object all the invariant features are extracted from object. According to this extracted features invariant feature vectors are generated. They are also relatively easy to extract which makes very

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

IJARCCE

*International Journal of Advanced Research in Computer and Communication Engineering*
*Vol. 4, Issue 4, April 2015*

suitable to match against large database. This calculated feature vectors are compared to vectors from database. It will get all the vectors from database and calculated feature vector is compared to vector by from database.

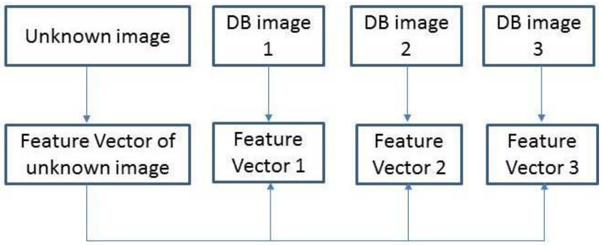


Fig. 2 Feature Extraction of unknown image

If there is any untested vector then it will calculate distance between them by using Euclidian distance algorithm. If calculated distance is smaller than previously calculated distance then it will change the PID of the best match and if there is no any untested vector and possible match is found then appropriate voice response is given to blind user.

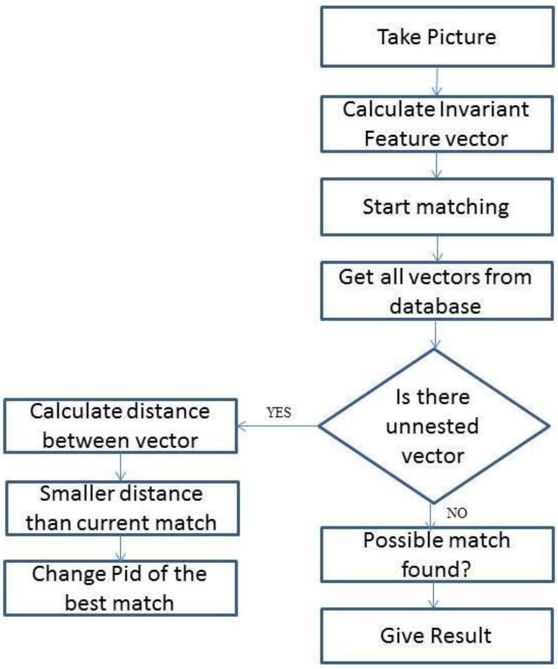


Fig. 3 Recognition ProcessPage Layout

## III. SERIALIZATION

Serialization is a mechanism of writing the state of an object into a byte stream. In computer science, especially in the context of Data Storage, Serialization is the process of translating data structures or object state into a format that can be stored either in a file or memory buffer, or it can be transmit across a network connection link. For eg. When the account holder tries to withdraw money from the server through ATM, the account hold with the withdrawal details will be serialized and sent to server where the details are used to perform operations. This will reduce the network calls as we are serializing sending to server and further request for information from client is not needed by the server. When the resulting series of bits is reread according to the serialization format, it can be used to create a semantically identical clone of the original object [7]. For many complex objects, such as those that make extensive use of references, this process is not straight forward. Serialization of object-oriented objects does not include any of their associated methods with which they were previously inseparably linked.

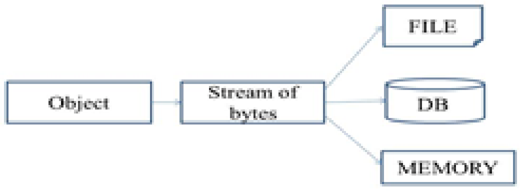


Fig.4 Serialization Process

## IV. NORMALIZATION

While capturing two dimensional images, based on the camera placement there are four possible basic forms of planar object shape distortions-rotation, scaling, translation, and skewing. A good shape descriptor should be invariant to these distortions. It is thus important to normalize the shape patterns in its original and various distorted forms, such as scaled, rotated, translated or skewed forms, so that they all, more or less, resemble similar to each other. When an appropriate set of features are extracted only after such normalization, shape classification yields much better accuracy. Normalization algorithm to normalize the shapes, called shape compacting, involves the following steps:

1. Computing the shape dispersion matrix M,
2. Aligning the coordinate axes with the Eigen vectors of M, and
3. Rescaling the axes using the eigenvalues of M.

### A. *Shape Distortion and Normalization*

In the normalization process, the alignment of the coordinate axis uses the dispersion matrix M and takes care of the rotation of the object. Rescaling the coordinate axes is integral component of shape compacting and it uses the dispersion matrix. The basic philosophy of shape normalization process is that after the normalization operation, the shape will have a dispersion matrix equal to an identity matrix multiplied by a constant. This is an indication that the shape is in its most compact form. To compute the dispersion matrix first we calculate the shape centroid.

$$\bar{x} = \frac{\sum x \sum_y x.f(x,y)}{\sum x \sum y\, f(x,y)} \quad ,\bar{y} = \frac{\sum x \sum_y y.f(x,y)}{\sum x \sum_y f(x,y)}$$

$$M = \begin{matrix} m1,1 & m1,2 \\ m2,1 & m2,2 \end{matrix}$$

Where

$$m_{1,1} = \left(\frac{\sum x \sum y x^2.f(x,y)}{\sum x \sum y\, f(x,y)}\right) - x^2,$$

$$m_{2,2} = \left(\frac{\sum x \sum y y^2.f(x,y)}{\sum x \sum y\, f(x,y)}\right) - y^2$$

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

*International Journal of Advanced Research in Computer and Communication Engineering*
*Vol. 4, Issue 4, April 2015*

$$m_{1,2} = m_{2,1}\left(\frac{\sum x \sum yx.y.f(x,y)}{\sum x \sum y\, f(x,y)}\right) - \bar{x}.\bar{y}$$

The shape dispersion matrix essentially performs the same function; it normalizes a shape by making it compact.

B. *Shifting And Rotating The Coordinate Axes*

The origin of the coordinate system is shifted to the center of the shape and then the coordinate system is rotated according to the eigenvectors of the dispersion matrix M. The matrix M has two eigenvectors El and E2 corresponding to the eign values A1 and X2. The two normalized eigenvectors El and E2 of M are computed as follows:

$$\alpha 1, \alpha 2 = \frac{m_{1,1} + m_{1,2} \pm \sqrt{\left(m_{1,1} + m_{1,2}\right)^2 + 4m_{1,2}^2}}{2}$$

And

$$E1 = \begin{bmatrix} e_{1x} \\ e_{1y} \end{bmatrix} = \begin{bmatrix} \dfrac{m_{1,2}}{\sqrt{(\alpha 1 - m_{1,1})^2 + m_{1,2}^2}} \\ \dfrac{\alpha 1 - m_{1,1}}{\sqrt{(\alpha 1 - m_{1,1})^2 + m_{1,2}^2}} \end{bmatrix}$$

$$E2 = \begin{bmatrix} e_{2x} \\ e_{2y} \end{bmatrix} = \begin{bmatrix} \dfrac{m_{1,2}}{\sqrt{(\alpha 2 - m_{1,1})^2 + m_{1,2}^2}} \\ \dfrac{\alpha 2 - m_{1,1}}{\sqrt{(\alpha 2 - m_{1,1})^2 + m_{1,2}^2}} \end{bmatrix}$$

Now we can construct a matrix R from El and E2 by

$$R = \begin{bmatrix} E_1^T \\ E_2^T \end{bmatrix} = \begin{bmatrix} e1,x & e1,y \\ e2,x & e2,y \end{bmatrix}$$

Since M is real and symmetric, El and E2 are orthogonal to each other. Furthermore, they are normalized to unit length. Thus, R is an orthonormal matrix. We now transform the coordinate system by first translating the origin to the shape center and then multiplying the coordinates with matrix R. Now, each object pixel location (x, y) will have a new location (x', y') given by:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = R.\begin{bmatrix} x - \bar{x} \\ y - \bar{y} \end{bmatrix}$$

Since R is an orthonormal matrix, the geometric interpretation of the trans- formation by R is pure coordinate rotation. The new coordinate axes are in the same directions as El and E2. The dispersion matrix D of the translated and rotated shape is given by:

$$D = \begin{bmatrix} \alpha 1 & 0 \\ 0 & \alpha 2 \end{bmatrix}$$

*C. Changing Scailes Of Bases*

In the previous step, we have rotated the coordinate system so that the new X-axis points in the direction in which the shape is most dispersed. The effect of the rotation on the dispersion matrix is that now it is a diagonal matrix. Since our objective is to have a shape whose dispersion matrix is a scaled identity matrix, in this last step we will change the scales of the two axes according to the eigenvalues XI and X2. That is, for an object pixel location (d. y'). the new location (z", y") is obtained through a transformation defined by W.

$$\begin{bmatrix} x^N \\ y^N \end{bmatrix} = W.\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} k/\sqrt{\alpha 1} & 0 \\ 0 & k/\sqrt{\alpha 2} \end{bmatrix}.\begin{bmatrix} x' \\ y' \end{bmatrix}$$

Where k is a system-wide constant. Since W is a diagonal matrix, the effect of the above step on the shape is to change the scales of the two coordinate basis vectors so that the shape is in its most compact form and with a normalized size. The result of translation, rotation, and scale invariance compact form generation has been shown in figure

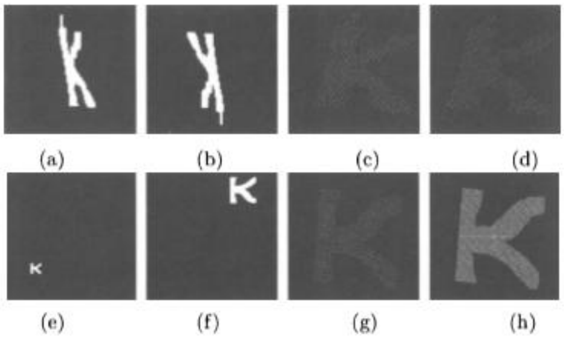


Fig.5 Normalized images by shape dispersion matrix: (a)-(b) rotated image of letter K, (c)-(d) normalized results, (e) -(f) scaled and translated version of letter K image (g)-(h) normalized results.

## V. CONCLUSION

Based on pilot studies conducted on a group of blind people, we concluded that, the object detection application is good for detection of objects in around a radius of 2 meters. It is affected by the lighting conditions in the region where the object is located. Proper alignment is essential while holding the device for better results.

## ACKNOWLEDGMENT

HANKS TO OUR PROJECT GUIDE **PROF. N.R. SONAWANE,** HEAD OF DEPARTMENT **PROF. G.V.GARJE** AND ALL STAFF MEMBERS OF INFORMATION TECHNOLOGY DEPT. FROM PVG'S COET, PUNE.

## REFERENCES

[1]. http://www.who.int/mediacentre/factsheets/fs282/en/
[2]. http://www.androidcentral.com/georgie-smartphone-visually-impaired-now-available-uk
[3]. Kane, S. K., Jayant, C., Wobbrock, J. O., and Ladner, R. E. Freedom to roam: A study of mobile device adoption and accessibility for people with visual and motor disabilities. In Proceedings of ASSETS 2009 (2009).
[4]. Voiceover,2009. http://www.apple.com/accessibility/voiceover/.
[5]. Kane, S. K., Bigham, J. P., and Wobbrock, J. O. Slide rule: making mobile touch screens accessible to blind people using multi-touch interaction techniques. In Proceedings of ASSETS 2008 (2008), 73–80.
[6]. http://www.tutorialspoint.com/java_dip/open_source_libraries.html
[7]. http://en.wikipedia.org/wiki/Serialization.