

ANDROID BASED OBJECT RECOGNITION AND MOTION DETECTION TO AID VISUALLY IMPAIRED

Neha Bari, Nilesh Kamble, Parnavi Tamhankar.

Padmabhooshan Vasantdada Patil Institute of Technology, Bavdhan, Pune(Pune University), India.

barineha29@gmail.com, nvkkamble93@gmail.com, tamhankarparnavi@gmail.com

ABSTRACT

In this paper we present the main features of software modules dedicated to the aid of visually impaired or blind users. The main aim of developing this software is to reduce or eliminate the need of separate dedicated devices for object recognition and motion detection. The software modules are designed for Android operating system, used in majority of the smartphones today. There are two main trainable (ANN based)modules, namely, the object recognition module and the motion detection module. Image processing algorithms used to identify the objects and detect motion are described. Notification to the users is given by means of verbal messages in this system.

KEYWORDS: Image Processing, Motion Detection, Object Recognition, Visually Impaired.

INTRODUCTION

The visually impaired and blind people face numerous challenges in routine tasks such as making coffee or crossing roads. Identification of objects and movement in the surroundings is a primary challenge for them which normal sighted people take for granted. We aim to provide some assistance to them via software based on android platform which will help ease the difficulty of performing the task of object identification and motion detection. There are about 285 million visually impaired people in the world, around 39 million of which are blind.

Currently, the most popular choice of smartphones among visually impaired users is iPhone or Android based phones. Commonly, the non-operating system devices are not preferred by blind users as they do not offer special functions such as text to speech conversion.

One can find a number of dedicated devices for navigation and object recognition. These wearable devices have the disadvantage that they are expensive in comparison to software. Also, the blind users are required to carry a number of gadgets and devices, each for a different purpose such as object identifiers, navigators and mobile phones.

Software that will run on the blind user's smartphones will eliminate the need to carry several devices.

Development of robust algorithms for object detection on Android is gaining popularity. Most algorithms focus on desktop applications.

The other solutions available at present include:

- Low-tech labelling systems in which labels are attached to objects, e.g. with tactile signs or text messages in Braille.[1]
- High-tech systems that employs 1-D and 2-D barcodes, talking labels or radio-frequency identification devices (RFID).RFID can be used to search for objects at short distances to which RFID tag was applied using an acoustic signal.[3]

These systems are not only costly but they also require continuous maintenance to keep them up to date.

In this paper we aim to present an aid for visually impaired and blind people for object recognition and motion detection. The software is based on image processing system running on Android platform. The object recognition and motion detection notifications are provided to the user via verbal messages.

The paper is structured as follows. In section I we review software applications for mobile devices that are specially devoted to blind users. In section II (A) image processing and analysis algorithms that were applied in object recognition module in the developed applications are described. In section II (B) image processing and analysis algorithms that were applied in motion detection module in the developed applications are described.

1. SECTION I

We have chosen the Android platform because of the immense popularity of Android devices. According to Gartner's analysis in the end of 2012 this mobile platform has gained 70% share of the smartphone market. [1] Further it also provides built in speech synthesis software.

Google search through voice has been available since initial release. Voice actions for calling, texting, navigation, etc. are supported on Android 2.2 onwards. As of Android 4.1, Google has expanded Voice Actions with ability to talk back and read answers from Google's Knowledge Graph when queried with specific commands. The voice search and TalkBack services in android facilitate use for blind people. Further, most smartphones have built-in, good quality digital cameras, GPS services, audio and video recorders, etc. These can be utilised by the dedicated software with ease.

A number of image analysis software's for mobile devices can be found in the market. The EyeRing project is one such application. It is a finger-worn device intended as an aid for blind users. This device communicates with an Android phone which runs the speech processing and image analysis algorithms. It provides facilities such as bank note recognition, colour recognition and distance calculation. It also works as a virtual cane.[5] However the major drawbacks of EyeRing are its high cost and that the blind user has to wear an additional device. Another software developed to aid blind and visually impaired people is Recognizer developed by LookTel.[4] This application is developed for iPhones and can recognize objects that are previously stored in a local database. In most cases it requires a sighted person to capture objects and an audio message is to be recorded identifying the objects. It is intended to help visually impaired people identify household objects. But, for images with arbitrary orientation of objects, the results of recognition are not specified [1]. Further, **Matusiak et al** provide software modules for object recognition, colour detection and light source detection. They have used a combination of SIFT and FAST algorithms. However, the modules do not have any provision for motion detection. Performance of the algorithms used depends on the quality of the built in camera.

2. SECTION II (A)

The application allows recognizing objects from images recorded by the camera of a mobile device. The object recognition algorithm should be insensitive to image registration parameters, i.e. scale, rotation and lighting conditions. Moreover, the recognised object should be robustly detected and localised in the image context (e.g. among other similar objects).

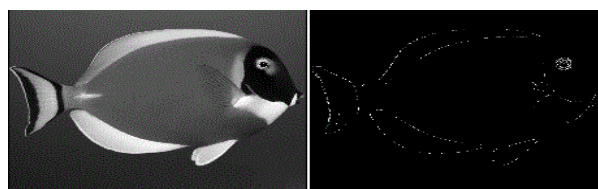
An RGB image is captured using the mobile phone camera. First, this image is blurred to reduce effect of noise. A 5X5 kernel is used for the same and average of the R, G and B values for the 25 pixels is found and applied to the central pixel. Blurring is done in order to reduce defects in the image such as small dark spots. This image is then converted to Grayscale format. Thus, now only the intensity

information of the image is contained in the pixel. This image which is made up of shades of grey is obtained by iteratively taking the average of the R, G and B values for each pixel forming the image. Fig3 represents the flow of image detection in our module. Following functions are performed.

i. Edge Detection

The Sobel operator is used for edge detection. We use a 3X3 Sobel matrix which is convoluted with our source image. The x-coordinate is defined here as increasing in the right direction changes, and the y-coordinate is defined as increasing in the down direction changes. According to this the gradient can be approximated. These gradient approximation in the x and y direction can be used to calculate Gradient Magnitude.

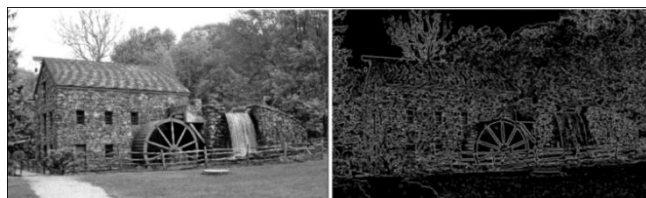
After calculating Gradient for the image, the required Edges can be detected as is seen in Fig1. The required features of the image (i.e. object to be detected) are extracted using thresholding method. In this method, a binary image is generated with only black and white colours. The required features are converted to white (or black) and the background is set to black (or white), respectively.



BeforeAfter
Figure1: Edge detection

ii. Thresholding

Thresholding is the simplest method of image segmentation. From a grayscale image, thresholding can be used to create binary images with only black or white colours. It is usually used for feature extraction where required features of image are converted to white and everything else to black (and vice-versa). Fig2 represents a thresholded image made up of only black and white colours.



BeforeAfter
Figure2: Thresholding

iii. Blob Detection

The next operation performed is blob detection. Blob detection refers to mathematical methods that are aimed at detecting regions in a digital image that differ in properties, such as brightness or colour, compared to areas surrounding those regions. Informally, a blob is a region of a digital image in which some properties are constant or vary within a prescribed range of values; all the points in a blob can be considered in some sense to be similar to each other. Since we have separated foreground (say black) and back ground (white) of an image in thresholding process. Hence by using those intensity values we compare x-coordinates of an image to find x min and x max. Similarly we do for Y-axis and find y min and y max and store it in a matrix and draw a blob over an object.

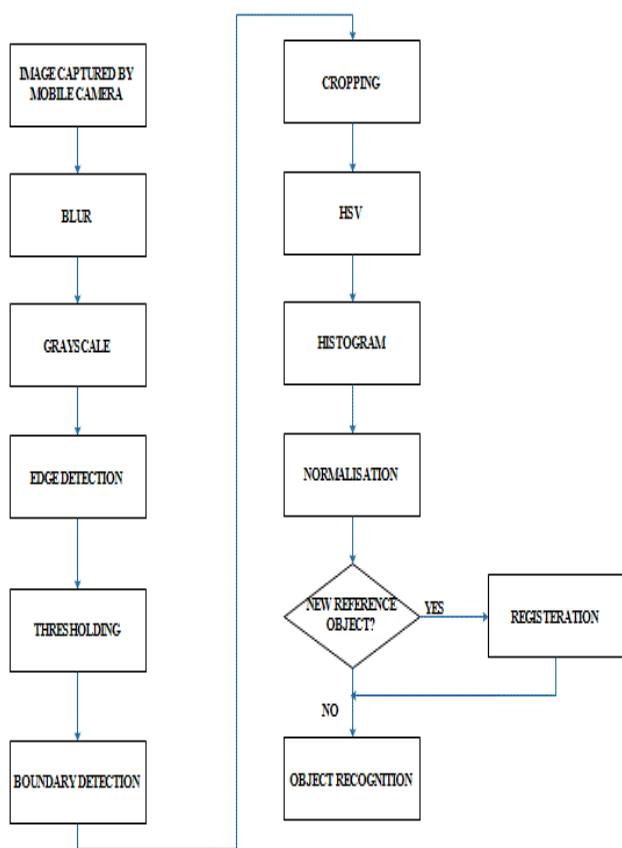


Figure3: Flow Diagram for Object Detection.

iv. Feature Detection

The SIFT (Scale-Invariant Feature Transform) was applied in the developed application. The SIFT is considered as a very power computer vision algorithm for detecting and describing local image features. SIFT allows to compute feature descriptors strongly independent on the image registration conditions. These

descriptors are further used to recognize objects in the proposed application.

Further, a key point detection procedure is performed. To improve performance of the application the Features from Accelerated Segment Test (FAST) algorithm was implemented. It is one of the fastest corner detection algorithms. The FAST corner detector is commonly used to track objects in different frames [2]. That is, FAST corner detector algorithms extract feature information, and rotation and movement information in different frames can be computed through feature matching, which is often based on a distance between the vectors, e.g. the Euclidean distance of feature vectors. In the application, corner information is extracted from the input image using the FAST corner detector and objects are recognized via BPNN (Back Propagation Neural Networks) machine learning. The number of interest points is then adjusted by normalization of extracted corners. [2] The FAST corner detector finds corners by examining a circle of sixteen pixels around the corner candidate. This candidate is detected as a corner if the intensities of a certain number of contiguous pixels are all above or all below the intensity of the center pixel by some threshold. The extracted interest points lie on distinctive, high-contrast regions of the image.

v. HSV conversion

The RGB colour images are converted into the HSV (Hue Saturation Value) colour images. This colour space enables to represent the colour in a single parameter, i.e. the H component, whereas the S component is the saturation parameter of the recognized colour.

vi. Histogram

Key points descriptor obtained from SIFT method is based on gradient magnitudes computed for 16 or 4 pixels adjacent to a key point. These values are used to form the histogram. By looking at the histogram for a specific image a viewer will be able to judge the entire tonal (i.e. representation of variation in the perception of a colour) variation of an image. The left side of the horizontal axis represents the black and dark areas, the middle represents medium grey and the right hand side represents white area. Thus if an image is dark then majority points will be located in the left side of the histogram similarly the grey points and the lighter points of an image can be located respectively in the middle and the right side of an histogram.

vii. Object Recognition

The BPNN is used to authenticate objects. A BPNN learns by example, that is, we must provide a learning set that consists of some input examples and the known-correct output for each case. The BPNN that uses the input-output examples to show the network what type of behaviour is expected has similarity with authentication in pattern recognition. After successful recognition we alert user using text to speech conversion function of android

3. SECTION II (B)

The second module of our application is the motion detection module. Fig4 shows the flow of operations performed in this module. Here a video stream is taken using camera of android mobile phone and is processed to estimate motion. The aim of motion detection is to identify moving objects from a sequence of image frames especially traffic which will alert the blind user. Several approaches have been proposed to the problem of motion segmentation. In conventional video surveillance systems, most of the motion detection is performed by using thresholding methods

A video stream is nothing but the stream of images taken continuously after 0.25 seconds. Hence we work on these images to detect motion and on successful detection we alert the user about it. Along similar lines, users are alerted when the motion (specifically moving vehicle) has stopped and there is no motion in front of the camera.

As seen in the object detection module, the image is blurred to reduce effect of noise. A 5X5 kernel is used for the same and average of the R, G and B values for the 25 pixels is found and applied to the central pixel. Blurring is done in order to reduce defects in the image such as small dark spots. This image is then converted to Grayscale format. Thus, now only the intensity information of the image is contained in the pixel. This image which is made up of shades of grey is obtained by iteratively taking the average of the R, G and B values for each pixel forming the image.

i. Background image registration

Before the moving objects can be identified, a background template must be built. The foreground cannot be removed so the ideal background image cannot be retrieved. But the moving objects do not exist in the same location in each image of a real-time video sequence. The grey values of pixels which have the same location in each frame of the video sequence are averaged to represent the grey value of the pixel located in the same place in the approximate background. An average value of pixels in the same location of each frame in a video sequence is calculated. To

simplify, the approximate background is also called “background template” [6]. These features are then stored in database to compare with newly captured images.

ii. Image compression

The objective of image compression is to reduce irrelevance and redundancy of the image data in order to be able to store or transmit data in an efficient form. Since we have to deal with a number of images during motion detection, compressing an image will ensure fast performance of proposed system

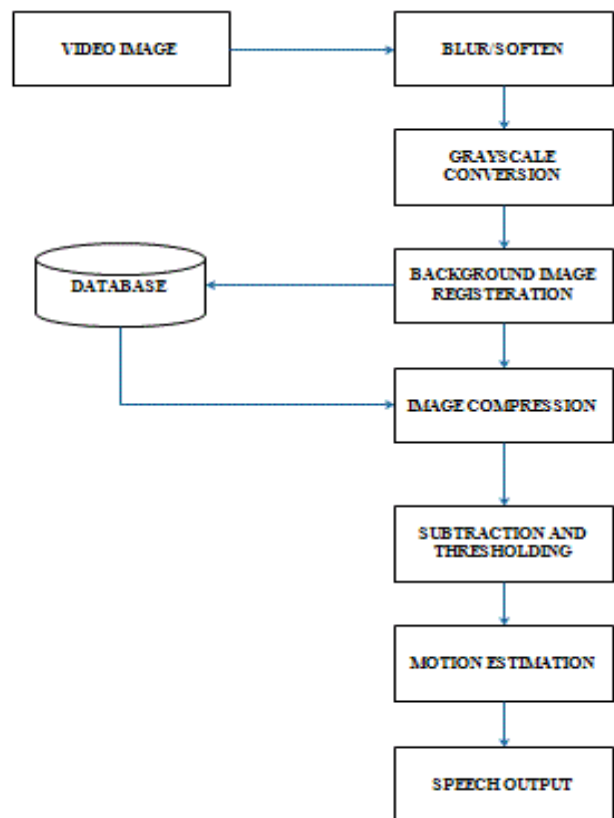


Figure4: Flow Diagram for Motion Detection.

iii. Subtraction and Thresholding

Here we separate foreground and background of an image in order to estimate motion. We set a threshold value to perform this separation. The threshold value is determined arbitrarily.

iv. Motion estimation

If there is drastic change in pixel values of previous loaded image and the pixel values of recent images, then this drastic change implies

presence of motion in front of the camera. Minor changes attributed to movement of the blind user and camera is neglected.

Framework, *IEEE Transactions On Intelligent Transportation Systems*, Vol. 13, No. 2, June 2012

v. **Speech Output**

This module has been added specifically to alert user about the motion in front of them and to convey the results of object detection to the blind users (especially when crossing roads). If motion is detected, an alarm is raised to alert the users. Due to this, blind users can take necessary precautions or stop for a while till the motion in front of them stops. After this again an alarm will be raised indicating that motion has stopped so that the users can proceed.

4. **CONCLUSION**

Image processing and recognition algorithms based on android platform and dedicated for blind users are proposed. Namely the object recognition algorithm and motion detection algorithm. This application will aid visually impaired and blind users in day to day activities. It will provide a portable solution to the problem of object and motion detection without the need to carry any additional devices dedicated to the same.

REFERENCES

- [1] K. Matusiak, P. Skulimowski And P. Strumillo, **Object recognition in a mobile phone application for visually impaired users**, IEEE Trans., 2014.
- [2] Kanghun Jeong, Hyeonjoon Moon, **Object Detection Using FAST Corner Detector Based On Smartphone Platforms**, 2011 First ACIS/JNU International Conference on Computers, Networks, Systems, and Industrial Engineering.
- [3] Alessandro Dionisi, Emilio Sardini, Mauro Serpelloni, **Wearable Object Detection System for the Blind**, IEEE 978-1-4577-1772-7/12, 2012
- [4] LookTel, <http://www.looktel.com/recognizer> Accessed 5th August 2014.
- [5] Nanayakkara S. C., Shilkrot R. And Maes P. **EyeRing: An Eye on a Finger**, Intl. Conf. Human Factors In Computing (Chi 2012).
- [6] Fan-Chieh Cheng and Shanq-Jang Ruan, **Accurate Motion Detection Using A Self-adaptive Background Matching**