BRAC UNIVERSITY

Inspiring Excellence

Staircase and Escalator Detection for Visually Impaired

A THESIS

Submitted to the School of Engineering and Computer Science, BRAC University, In partial fulfilment of the requirements for the degree of Master of Science in Computer Science and Engineering

By

Aniqua Nusrat Zereen

CERTIFICATE OF APPROVAL

The thesis titled "Staircase and Escalator Detection for Visually Impaired" is completed under my supervision, meets acceptable presentation standard and can be submitted for partial fulfilment of the requirement for the degree of Master of Science in Computer Science and Engineering from the department of Computer Science & Engineering, BRAC University.

Dr. Jia Uddin

Assistant Professor

Department of Computer Science & Engineering

BRAC University

DECLARATION

I hereby certify that this material, which I now submit for assessment on the programme Master of Science in Computer Science and Engineering is entirely my own work, that I have exercised reasonable care to ensure that the work is original, does not to the best of my knowledge breach any law of copyright and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged within the text of my work.

Aniqua Nusrat Zereen

ID number: 14366004

ACKNOWLEDGEMENTS

All praises are due to the merciful Almighty Allah, the most exalted, the most Beneficent, the most Merciful who blessed me to be here at BRAC University for pursuing the master degree.

My journey towards the Master degree would not have been possible without the help of many people. It is my great pleasure to take this opportunity to thank them for the support and advice that I received.

My deepest gratitude should go to my Supervisor, Dr. Jia Uddin, PhD, Department of Computer Science and Engineering. His comprehensive and constructive comments, his significant support throughout this work gave me strength during the work. I attribute the level of my Master's degree to his encouragement and effort and without him this thesis, too, would not have been completed. One simply could not wish for a better or friendlier supervisor. I would like to thank to my co-supervisor Sonia Corraya, for the continuous support of my master study and related research, for her patience, motivation and immense knowledge. Her guidance helped me in all the time. I owe my most sincere gratitude to our Chair of the Department of Computer Science and Engineering, Prof. Dr. Md. Haider Ali for his important support for completing this degree.

I would like to thank my parents for inspiring me in every step of my life. Supporting me spiritually and providing me the best education. My whole family who gave me strength to complete this degree. My husband for his unconditional support and trust on me. This thesis is dedicated to them.

To my husband, sister and parents

CONTENTS

List of Tables

LIST OF FIGURES

ABSTRACT

In this thesis, two different methods are presented for staircase detection. First method works with real time captured still image and second method with real time video. Stairway detection and identification of up stair and down stair is important for both independent and safe navigation of visually impaired. Proposed first method for detecting up stair and down stair using Gabor filter and Support Vector Machine (SVM). Energy distribution as a feature has been extracted from 40 filtered images using Gabor filters with 5 scales and 8 orientations. These features are trained and tested with four different classes, up stair, down stair, freeway and other. The overall classification accuracy of the proposed method is 92.9% based on experimental result. In second method, real time video is captured with Microsoft Kinect. Detection of real time moving object along with the moving direction in respect with visually impaired people is a challenging research area. The recent advancement in technology for real world scene capturing and portable devices like Microsoft Kinect necessitate the need of more robust and faster techniques for assisting blind navigation. The objective of this study is to develop a suitable and an effective technique for moving object detection along with its moving direction in indoor environment. Depth information of the font scene of a blind people is captured using Microsoft Kinect version 1. Three consecutive depth frames are extracted from video taken in one second and Distance Along Line Profile graph is generated for four predefined lines of each depth frame. These line profile graphs are analyzed for detecting presence of any moving object and the moving direction. After detailed investigation, experimental result shows that the proposed method can successfully detect moving object along with its direction and still objects with 92% and 87% accuracy respectively. The overall accuracy of the proposed second method is 90%.

CHAPTER I

INTRODUCTION AND OVERVIEW

1.1 MOTIVATIONS

285 million people with visual impairments have been reported by The World Health Organization, among them 39 million are totally blind and 246 million having low vision [1]. Understanding front scene is the fundamental test for blind individuals. To sense their surroundings, whether any obstacles are there or moving towards them, they need to depend on guide cane or to touch physically. Recently technologies are getting advanced for portable devices like camera, Microsoft Kinect [2] which necessitates the need of more robust and faster techniques for assisting blind navigation. Object identification, detection, classification, moving object tracking, obstacle freeway detection, staircase detection and crossway detection are some recent developments for blind assistance [29, 31, 33, 34 and 35]. Among all these identifications types, staircase detection is one of the most essential and important one for blind navigation. Existing stair detection systems are mostly based on sensor, stereo, sonar, laser scanning, Microsoft Kinect, cell phone's camera etc. [14, 15, 16, 18, 19, 20, 21, 22]. Some common limitations of using those methods are the accuracy, light source requirement, requirement of multiple hardware setups on blind person as well as in the whole environment, complex or too much informative output, longer processing time, poor performance in noisy (image or sound) environments, costing, robustness, manual update requirement and so on.

In this thesis, two different methods are presented for staircase detection. First method works with real time captured still image and second method with real time video. Proposed first method is based on still images taken from camera which is convolved with 2D Gabor filter and Support Vector Machine (SVM). As a feature, Energy distribution of 40 Gabor filtered images with 5 scale and 8 orientations have been taken and fed to SVM to classify into 4 classes naming, up stair, down stair, freeway and other. For the purpose of training and testing with SVM, different images have been taken in account. In second method, real time video is captured with Microsoft Kinect.

The objective of this proposed method is to develop a suitable and an effective technique for moving object detection along with its moving direction in indoor environment. These objects include staircase, escalator and others. From the depth information of the captured one second video, 3 consecutive depth frames are extracted and Distance Along Line Profile (DAP) graph is generated for 4 predefined lines of each depth image. These line profile graphs are analyzed for detecting presence of 6 types of categories, Freeway, Wall or door, Staircase, Other object-Still, Escalator-Moving and Other object-Moving.

1.2 THESIS CONTRIBUTION

The main contributions of this thesis are as follows.

- Gabor filter with scale 5, 8 orientations and SVM are used to classify still images taken by camera into 4 different classes naming, up stair, down stair, freeway and others. Features have been extracted based on energy distribution of filtered images. For training SVM, A total of 815 images of which 227 upstairs, 221 downstairs, 175 freeway and 192 others have been taken. For testing, real time input image is captured with camera. In addition, to test the performance of this system, different scale and orientations' $n\times m$ Gabor filtered images have been taken individually in account (e.g. 1×1, 1×2, 1×3… up to 7×8) where n stands for scale and m for orientations. Another 403 images of which 124 upstairs, 118 downstairs, 72 freeway and 89 others have been taken and tested with another 2 different classifiers along with SVM, Naive Bayes and Back Propagation Neural Network. Under SVM classifier different kernel functions, Pearson VII function, Radial Basis Function (RBF), Normalized Polynomial and Polynomial also were used to evaluate the performance of the system. From all the testing results, SVM with RBF kernel outperforms with 92.9%.
- In addition to detect staircases as up stair and down stair, freeway and other, second proposed method can also detect moving escalators along with its moving direction (upward or downward), other object's moving direction and wall or door with respect to the blind person's view point. Depth information of the font scene of a blind person is captured using Microsoft Kinect version 1 sensor.

From one second video, 3 consecutive depth images are extracted and Distance Along Line Profile graphs are generated for 4 predefined lines of each depth image. These line profile graphs are analyzed for detecting presence of any moving object and its moving direction. Proposed method can successfully detect moving object along with its direction and still objects with 92% and 87% accuracy respectively. The overall accuracy of the proposed second method is 90%.

1.3 THESIS OUTLINE

The rest of this thesis is summarized below.

- Chapter III represents Gabor filter and SVM based Stair case detection along with freeway and others type of object.
- Chapter IV represents depth image based real time object detection along with the moving direction method using Microsoft Kinect.
- Chapter V concludes this thesis.

CHAPTER II

BACKGROUND

2.1 INTRODUCTION

Navigation systems are being used widely nowadays to assist blind people moving freely and independently. Researchers has shown remarkable performance of different kind of navigation systems using sterao camera, sensors, laser scanning, camera, sonar, microsoft kinect etc. Recently image processing techniques are being widely used with microsoft kinect and other cameras for these assistive systems. Therefor in this chapter core knowledge of image processing and its outcomes using different techniques are reviewed.

2.2 IMAGE PREPROCESSING

2.2.1 RGB IMAGE AND RESOLUTION

RGB color model stands for red, green and blue. These colors are added together in various ways to produce a broad array in electronic systems to be displayed. A RGB image is taken using camera and converted into gray scale (0~255) for computational low cost. Each image are resized to 640×480 pixel resolution [3]. Further for detecting edges binary image of input image has been used. Figure 1 shows a) RGB image and b) converted gray image.

(a) (b)

Figure 1. Down stair image in (a) RGB, (b) Gray level.

2.2.2 DEPTH IMAGE

Depth image contains information of distance of the surfaces of object in the scene from a viewpoint. It is used in various detection methods e.g. object detection, distance measure, human detection, hand tracking [5,6,7,8]. Depth image can be taken using depth sensor cameras and convert it to gray level depth image. In image the darker diplayed portion indicates object is far away and lighter display portion indicates near. Figure 2 shows depth image and its darker and lighter portion. Resolution is resized to 640×480 for proposed system.

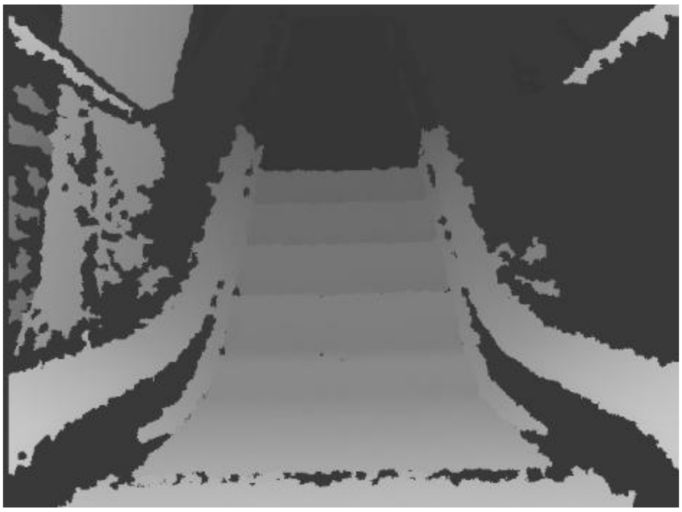


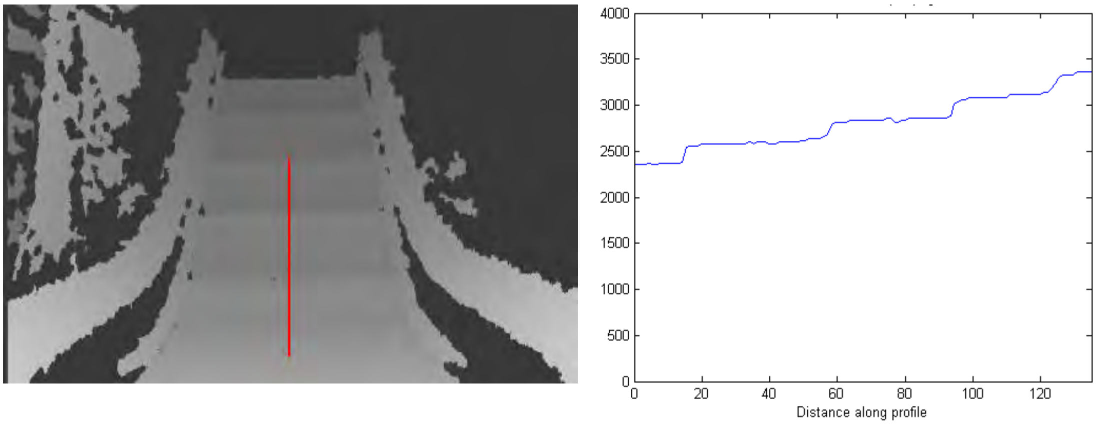
Figure 2. Depth image of escalator.

2.2.3 MORPHOLOGICAL EROSION AND DILATION

Morphological image processing is a technique which is a set of operators that can transform digital images to analyze the geometrical structures characterized as connectivity, shape, convexity, size, etc. There are four kinds of operators such as erosion, dialation , opening and closing. In this thsis erosion and dialation has been used to analyze depth images accordingly [52].

2.3 LINE PROFILE

Line profile of an image is the set of intensity values across the line that has been created with regularly spaced points [47]. Researchers has used this techniques to measure pixels intensity along multiple lines to get specific patterns which can be detected as specific objects [29].



(a) (b)

Figure 3. (a) Shows line profile colored red on escalator, (b) Pixel intensity along line profile.

2.4 Sobel Edge Detector

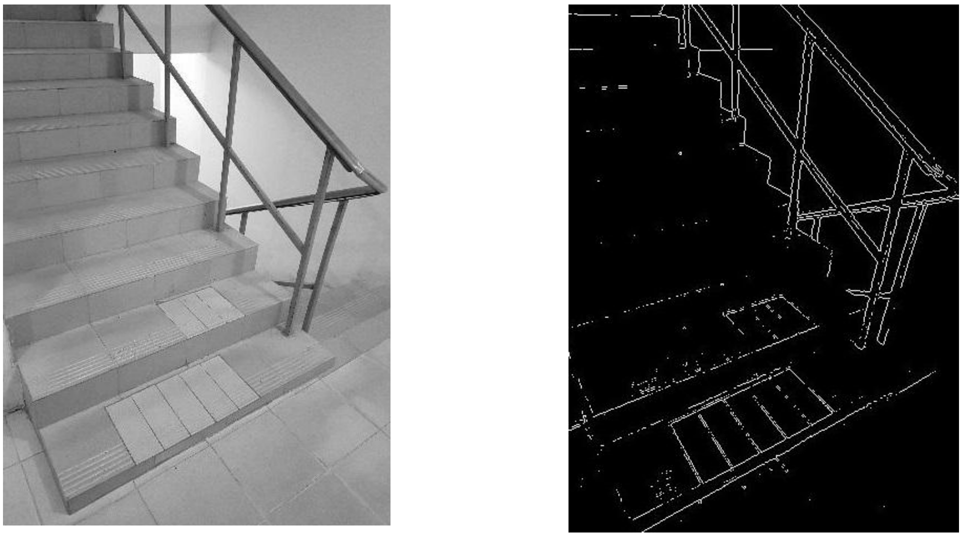
In image processing Sobel edge detector is commonly used for its better performance. Different kinds of object detection, surveillance system, pattern recognition has been done using this. In [48] author has showed a comparison between different edge detectors to analyze night vision where sobel outperforms. It has a operator which uses two 3×3 kernels which are convolved with the input image to compute approximations of the derivatives, one for horizontal changes and other one for vertical. The equations are below [51]:

$$G_x = \begin{matrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{matrix} * \mathrm{I} \tag{1}$$

$$G_y = \begin{matrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ +1 & +2 & +1 \end{matrix} * \mathrm{I} \tag{2}$$

$$G = \sqrt{{G_x}^2 + {G_y}^2} \tag{3}$$

Equation (1) and (2) represents G_x and G_y which are two images, each point contain the horizontal and vertical derivative approximations respectively. * denotes the 2-dimensional signal processing convolution operation. At each point in the image, the gradient approximations can be combined to give the gradient magnitude, using equation (3). Above equations are used to create the edge detected image in Figure 4.



(a) (b)

Figure 4. Up Stair image in (a) Gray level, (b) Edge detected by Sobel.

2.5 MICROSOFT KINECT

Microsoft Kinect sensor is affordable and has the ability to work in low light environments. By recognizing optical imprints using kinect, Zöllner et al. proposed system can guide a blind person [4]. The sensor is used to determine the distance from the user to objects from a view point. Mapped depth image is fed via wireless to a haptic glove [49]. Depth images are taken with the Kinect sensor, which comprises of a depth sensor and an RGB camera in Figure 5.

Figure 5. Microsoft Kinect.

The depth information represents distance from the sensor to the object. Lighter gray level in image represents object is near and darker represents farway.

2.6 Gabor Filter

This function is appropriate for a specific spatial location to differentiate between the objects of an image. Gabor filter represents frequency and orientation of detected edges in image. 2D Gabor filter is a Gaussian kernel function transformed by a sinusoidal plane. Gabor filters are associated with Gabor wavelets, which are designed for a number of dilations and rotations [12, 23, 25].

2.7 Support Vector Machine (SVM)

It is a statistical knowledge based classification system which separates the d-dimensional data into multi classes by finding hyper plane. For multi class data set a kernel function can be used to define multiple of non-linear relations between the data sets. There exists non-linear kernel functions which are polynomial, Gaussian radial basis function (RBF) and hyperbolic tangent. In this thesis RBF kernel is used due to its simplicity and dynamic non-linear classification capacity [26].

2.8 Chapter Synopsis

For better understanding all the techniques have been used in this thesis were explained briefly. Example images and equations to compute this technical terms are given in this chapter. Image or video taken for this purpose were preprocessed using morphological erosion and dilation, later using sobel edge detector, line profile, Gabor filter and SVM for different purposes, images have been categorized.

CHAPTER III

A NEW METHOD FOR UPSTAIR AND DOWNSTAIR DETECTION FOR VISUALLY IMPAIRED USING GABOR FILTER AND SVM

3.1 INTRODUCTION

Staircase detection and identification of up stair and down stair is important for both independent and safe navigation of visually impaired. Among various object detection and tracking, stairway detection is the most important challenge for ensuring secure mobility of blind persons. Image noise, feature selection, determining pre and post image processing technology, adaptive system performance are some of the common problems for facing this challenge

There exist some systems for assisting blind in staircase detection but lacks in advance technology use, accuracy, response time, user-friendliness and cost-effectiveness. Different image features (e.g. edge, intensity, texture) are generally considered for image understanding and staircase detection. Existing stair detection systems are mostly based on stereo, sensor, sonar, laser scanning, Microsoft Kinect, cell phone's camera etc. [14, 15, 16, 18, 19, 20, 21, 22]. In recent years, Gabor feature uses in face detection, fingerprint and texture segmentation [23-25] are appreciable and produces very good result. For texture based staircase detection, Gabor filter also responses remarkably [18]. There are more important and efficient features in Gabor feature vector to research and analyze for staircase detection and identification beyond Gabor texture feature. SVM is known as a good classifier and also performed well with Gabor filter in [17]. It also showed significant result using ultrasonic signal for detecting staircase [28].

Staircase has been recognized by many different techniques. Recently, researchers focus on representing visual information into high level interpreted information before sending to the visually impaired person. Using RGBD image, Munoz et.al have proposed a technique where parallel lines are extracted from RGB frames by Hough transform and depth frames are used to recognize upstairs, downstairs and negatives. These three categories are classified with Support Vector Machine (SVM) [9]. But this technique has considered only Gabor texture feature and performs better for down stair and negative category compared to up stair detection.

In paper [13], the authors modeled a stair's tread and rise through using cloud points, scene segmentation and geometry constraints. A Microsoft Kinect based system has poor performance in sunlight. So, this system is suitable for indoor light environment. Se and Brady's system used a technique taking into account Gabor filters to identify distant staircases. When a user approaches close to the staircase then system determines parallel edges, where convex and concave lines are divided using intensity variation. Staircase posture is evaluated by a homography search [10]. Hernandez et al. proposed a system which localizes stairways. This technique investigates the edges of stairs by applying planar motion tracking and directional filters. It recognized the horizontal edges by using Gabor filters. From the predefined set of horizontal edge segments, they extracted a hypothetical set by using a correlation method. Finally, discrimination method is applied to find the ground plane based on behavioral distance measurement [18]. These techniques did not recognized up stair or down stair separately. Based on Gabor filters and fuzzy fusion phase grouping (FFPG) system, Zhonget al. presented robot autonomous stairway climbing, stairway modeling in mapping and building reconstruction. Blur images captured by robot because of vibration and poor illuminating condition, has been filtered with Gabor filter to extract edges of staircase efficiently [11]. Position of vehicle on stairs and the orientation angle to stairs is estimated in [17]. This system used Gabor filter to extract stair edges for only upstairs.

To overcome the limitation of state-of-art models, we proposed a cost effective method using Gabor filters and Support Vector Machine (SVM) taking into account energy distribution as features derived from Gabor filtered images with 5 scales and 8 orientation which classifies staircases into 4 different classes named Upstairs, Downstairs, Freeway and Other.

3.2 PROPOSED METHOD

Overall workflow of the proposed method is presented in Figure 6. The functionality of the different blocks related to our method including input, Gabor filtering, classification and output details are presented in different subsections.

3.2.1 THE INPUT

Capture 2D image of the front screen with simple digital camera and convert the RGB images into Gray scale (0~255). Later, resize image with resolution setting to 640×480.

3.2.2 GABOR FILTERING

This function is appropriate for a specific spatial location to differentiate between the objects of an image. Gabor filter represents frequency and orientation of detected edges in image. 2D Gabor filter is a Gaussian kernel function transformed by a sinusoidal plane. Gabor filters are associated with Gabor wavelets, which are designed for a number of dilations and rotations [12, 23, 25]. Following equations are used to create different types of Gabor filtered features.

Complex

$$g(x, y; \boldsymbol{\lambda}, \boldsymbol{\theta}, \Psi, \sigma, \tau) = \exp\left(-\frac{x^2 + \tau^2 y^2}{2\sigma^2}\right) \exp(i(2\pi \frac{x'}{\boldsymbol{\lambda}} + \Psi)) \tag{4}$$

Real

$$g(x, y; \boldsymbol{\lambda}, \boldsymbol{\theta}, \Psi, \sigma, \tau) = \exp\left(-\frac{x^2 + \tau^2 y^2}{2\sigma^2}\right) \cos(i(2\pi \frac{x'}{\boldsymbol{\lambda}} + \Psi)) \tag{5}$$

Imaginary

$$g(x, y; \boldsymbol{\lambda}, \boldsymbol{\theta}, \Psi, \sigma, \tau) = \exp\left(-\frac{x^2 + \tau^2 y^2}{2\sigma^2}\right) \sin(i(2\pi \frac{x'}{\boldsymbol{\lambda}} + \Psi)) \tag{6}$$

Energy

$$E_{\lambda,\sigma,\theta}(x, y) = \sqrt{R^2_{\lambda,\sigma,\theta,0}(x, y) + R^2_{\lambda,\sigma,\theta,-\frac{\pi}{2}}(x, y)} \tag{7}$$

Where x'= $x$cos$\boldsymbol{\theta}$ + $y$sin$\boldsymbol{\theta}$ and y'= -$x$sin$\boldsymbol{\theta}$ + $y$cos$\boldsymbol{\theta}$, $\boldsymbol{\lambda}$ is the wavelength of sinusoidal factor, θ is the orientation, Ψ is the phase offset, σ is the standard deviation, τ is the spatial aspect ratio.

A filter bank has been created with the Gabor filters equation (4) consisted of different scales and orientation combinations. Based on trial and error, 5 scales and 8 orientation filters has been confirmed for the system after detail investigation with different scales and orientations (1×1, 1×2, 1×3… up to 7×8). All 40(5x8) filtered images are convolved with 2D Gaussian convolution. Later, energy distribution as features is extracted using equation (7) from these images.

3.2.3 CLASSIFICATION

For the proposed system, multi-object classification of support vector machine has been used [9]. It is a statistical knowledge based classification system which separates the d-dimensional data into multi classes by finding hyper plane. SVM has two phases in classification process, one is training and other one is testing. In this step SVM with Radial Basis Function (RBF) kernel [27] is used to train the system before with sample images of 4 different classes, Upstairs, Downstairs, freeway and others.

A total of 815 images of which 227 upstairs, 221 downstairs, 175 freeway and 192 others have been taken for training.

3.2.4 OUTPUT

Based on the classification result the final output is produced which could then interpreted as voice, vibration or any other means for the visually impaired person.

3.3 EXPERIMENTAL RESULT AND ANALYSIS

In order to evaluate the performance of proposed method we used MATLAB 2016 on a system having 4 GB RAM and 2.6 GHz Intel i-5 processor. The proposed system uses Gabor filter and support vector machine (SVM).

The new method presented in this chapter is able to work with indoor environment. Different kind of images of upstairs, downstairs, household accessories, wall, door and obstacle free front scenes are considered for this research work. Data are collected from open access data storage of internet and also captured with camera [46, 50].

In addition to evaluate the performance of the proposed method, Gabor filter with different scales and orientations have been taken. Filter combinations are (1×1, 1×2, 1×3… up to 7×8) where first value indicates scale and last value indicates orientation (e.g. 1×1). From these individual set of filters, energy distribution has been measured.

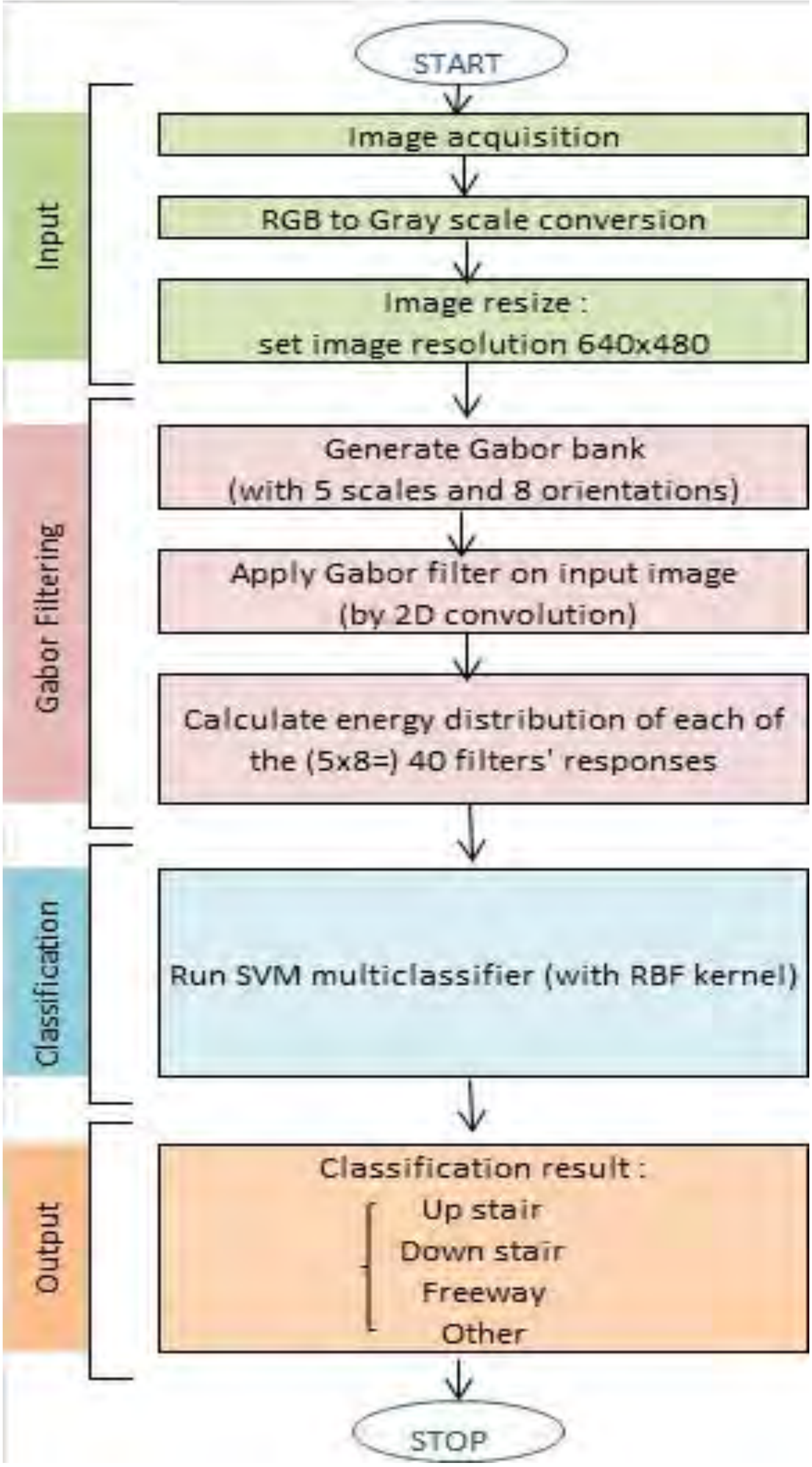


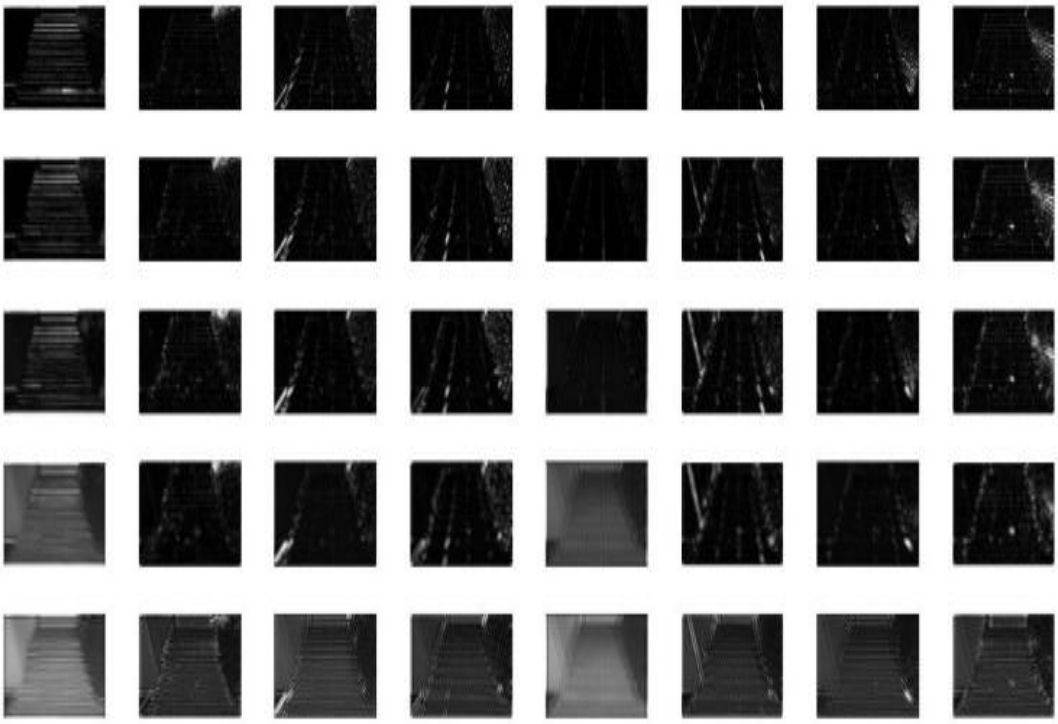
Figure 6. Flowchart of the proposed system for up stair and down stair detection.

For training purpose different kind of environment such as foggy, blur, low light, sunlight with different angle of view point have been taken in account. A total of 815 images of which 227 upstairs, 221 downstairs, 175 freeway and 192 others have been taken for training SVM. Another 403 images of which 124 upstairs, 118 downstairs, 72 freeway and 89 others have been taken for testing purpose.

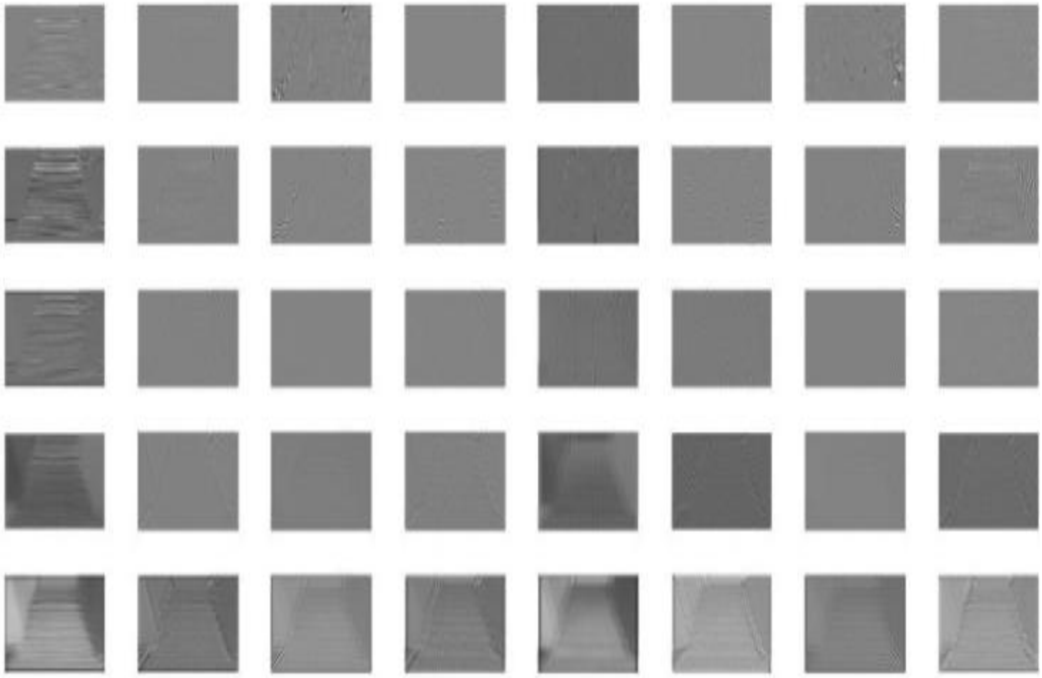
The resolution of captured RGB images are set to 640×480 pixels. Resolution varies for collected images from internet which are analyzed, respectively.

In the training phase, SVM is being trained with extracted features of training images with their corresponding classes. In the testing phase, real time input or images from testing set gets classified which is the actual classification process of SVM. In this work, SVM has 4 different classes named Upstairs, Downstairs, Freeway and Other. In addition, to test the dataset, other 4 features including Mean amplitude, Absolute values, Real part and Imaginary part of Gabor function using equation (4, 5, 6) are also considered. Figure 7 illustrates 3 different features (absolute, real and imaginary) of a sample input image. The feature vector size varies depending on scale and orientation combination. Later each of the feature type with all possible combination of Gabor filters individually has been trained and tested with SVM classifier. To compare result of SVM classifier, additional two classifiers, such as Naive Bayes and Back Propagation Neural Network (BPNN) [26] has been taken. Under SVM classifier, dataset has been tested considering different kernel functions Pearson VII function, Radial Basis Function (RBF), Normalized Polynomial and Polynomial to get the optimum result [27].

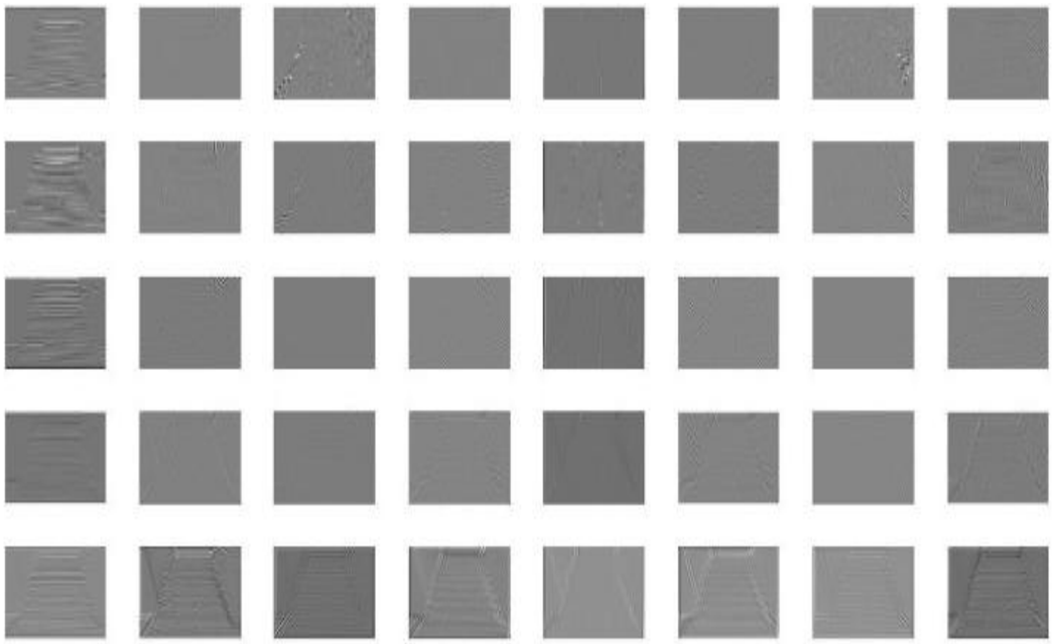
(a)



(b)



(c)



(d)

Figure 7. (a) Input image, (b) Absolute features, (c) Real features, (d) Imaginary features.

Table 1 shows the best test result of 403 images from each classifier based on feature type. After considering trial and error with all the scale and orientation combinations of 403 test images, system outputs best with 5×8 filters for each type features. Each classifier has been trained with all possible scale and orientation filters of 815 train images individually before testing. Considering the best kernel function, RBF for SVM result shows 92% accuracy with energy feature where other features show 82-89%. In addition, experimental results show (as depicted in Table 1, Figure 8 and 9), SVM classifier outperform than other classifiers such as Naive Bayes and BPNN. Naive classifier shows 84, 80, 70, 73, 71% classification accuracy where BPNN shows 88, 81, 72, 76 and 79%, respectively for energy, mean amplitude, absolute, real and imaginary features. It is important to note that significant result has been encountered by energy feature with 88 percent on average for all the classifiers showed in Table 1 and Figure 9.

Table 1. Gabor feature accuracy for three classifiers

Classifier	**Gabor Feature Accuracy (Scale × Orientation)***					Average
	Energy	***Mean Amplitude***	***Absolute***	***Real***	***Imaginary***	
SVM*	92	89	82	85	87	87
Naive Bayes	84	80	70	73	71	75.6
BPNN	88	81	72	76	79	79.2
Average	88	83.3	74.7	78	79	

* Best performed kernel for SVM is considered

* Best performed Gabor Scale × Orientation is considered

Table 2 shows testing accuracy of 403 images, choosing the best feature, energy and classifier, SVM with RBF kernel. Figure 10 depicted sample images of dataset including up stair, down stair, freeway and other. Among 124 upstairs 115 have been detected correctly, misclassified 7 as downstairs and 2 as other which shows 92.7% accuracy. 107 downstairs were detected correctly from 118 images, 5 as upstairs, 4 as other and 2 as freeway showing 90.7%. 66 out of 72 freeway images were correct, 6 were misclassified as other resulting 91.7%.

86 of 89 other images were correct, 1 as up stair and 2 as downstairs detected outputting 96.6% accuracy. Overall system performance shows 92.9% accuracy. Negative results can be minimized if richer dataset is used. Misclassification has been occurred because of low light and different object patters of other class's image.

Table 2. System performance result by test images considering energy with SVM

Test Image	Accuracy gained with Energy features (%)				
	Up stair	***Down stair***	***Freeway***	***Other***	**Total Image**
Up stair	115	7	0	2	124
Down stair	5	107	2	4	118
Freeway	0	0	66	6	72
Other	1	2	0	86	89
Accuracy	92.7	90.7	91.7	96.6	
Overall Accuracy			92.9		

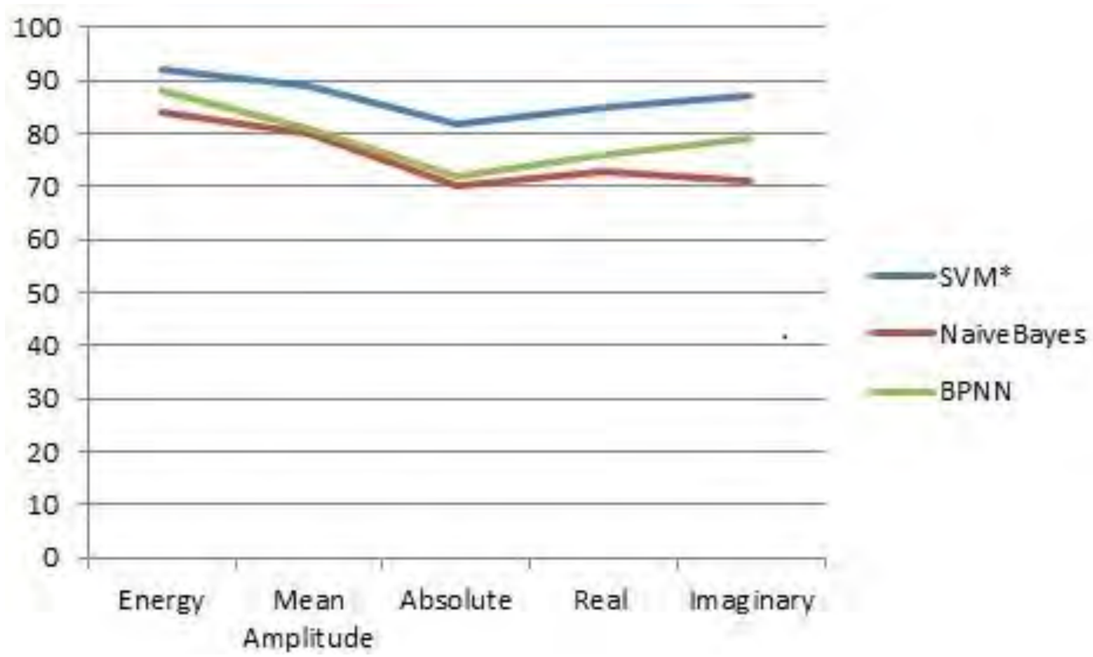


Figure 8. Accuracy of different classifiers for different statistical parameters.

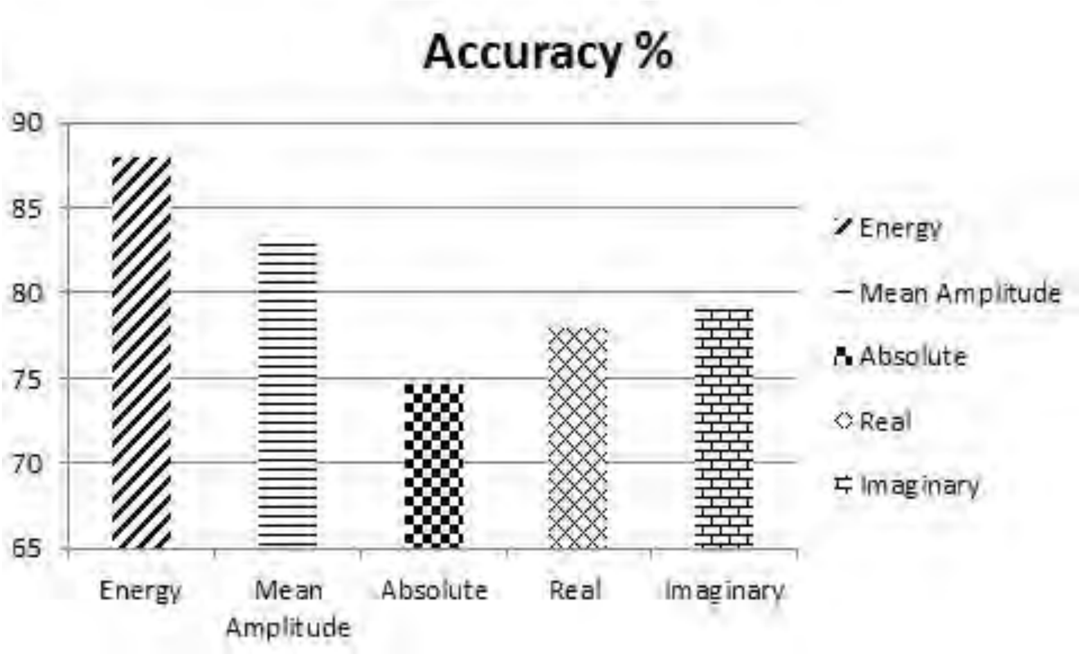
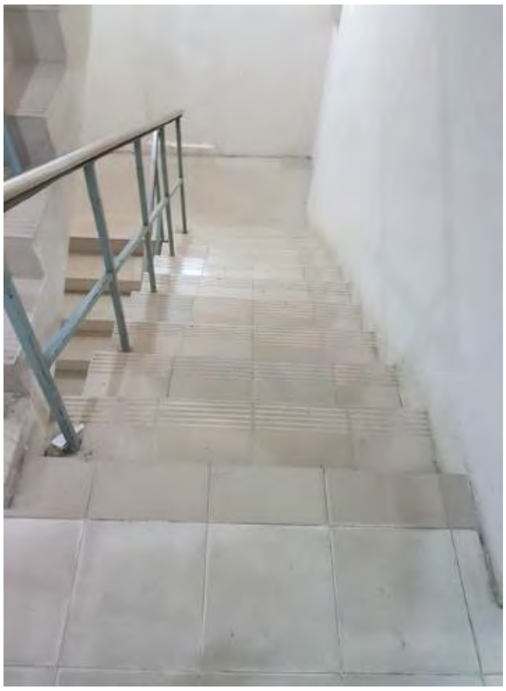
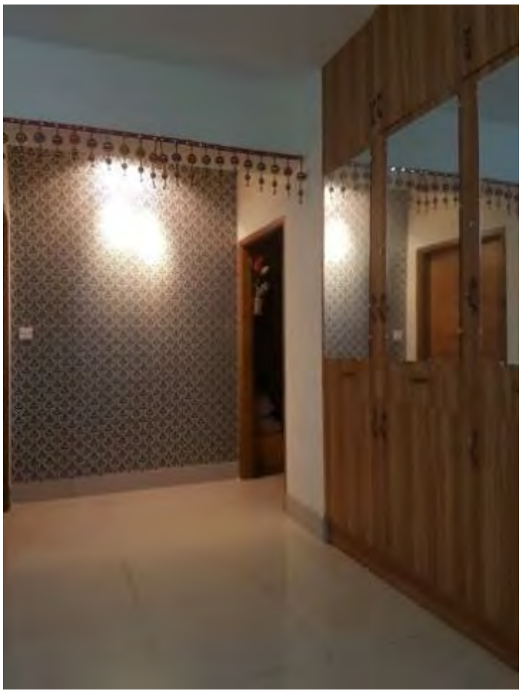


Figure 9. Average accuracy of different features based on three classifiers.

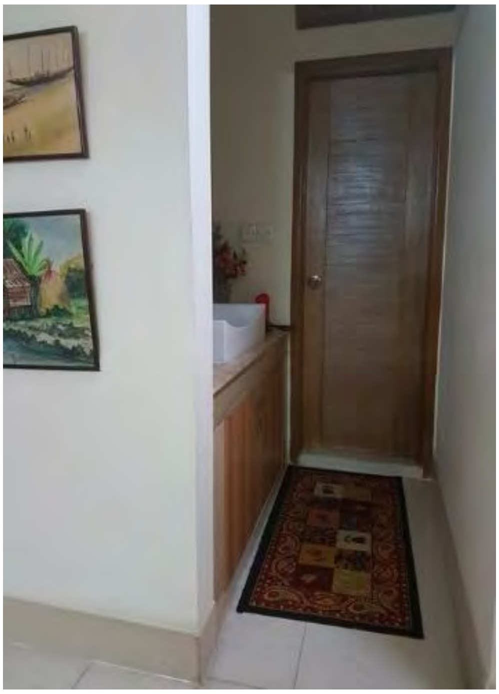
(a)



(b)



(c)



(d)

Figure 10. Sample image of (a) Up stair, (b) Down stair, (c) Freeway, (d) Other.

3.4 CHAPTER SYNOPSIS

In this chapter, a new method is presented for up stair and down stair detection for visually impaired using Gabor filter and SVM. Research focus was to identify the best Gabor feature and classifier for up stair, down stair, freeway and other object type detection. Five types of features (Energy distribution, Mean amplitude, Absolute, Real and Imaginary part) were trained individually with three different classifiers (Support Vector Machine (SVM), Naive Bayes and Back Propagation Neural Network (BPNN)). Energy features with SVM resulted the best performance with 92.9% accuracy.

CHAPTER IV

A NEW METHOD TO DETECT REAL-TIME OBJECT ALONG WITH THE MOVING DIRECTION FOR VISUALLY IMPAIRED PEOPLE

4.1 INTRODUCTION

Understanding front scene is the main challenge for visually impaired people. Blind people generally have to rely on guide cane or to touch physically to sense obstacle in their moving ways which is risky. The risk is higher when any object is moving towards the blind people from any direction. The main goal of this work is to make blind people aware of his/her environment and feel secure, independent while moving. Day by day with the advancement of technology, ways of blind navigation support increase as well as the number of blind people [44]. The recent advancement in technology for real world scene capturing and portable devices like Microsoft Kinect necessitate the need of more robust and faster techniques for assisting blind navigation. Object identification, object detection, object classification, moving object tracking, obstacle freeway detection, staircase detection, crossway detection all these are some recent developments for blind assistance. With the help of computer vision advancement, there are many existing methods for the above purposes. Traditional methods use video sequence, multiple sensors at different location, rotating or moving cameras, laser ray, sound echo based sonar system etc. to fulfill these purposes. Only a few works have considered a single sensing point which is important from blind person's view point. Front scene selection control should be in the moving person's hand. Some common limitations of using those methods are the accuracy, light source requirement, requirement of multiple hardware setups on blind person as well as in the whole environment, complex or too much informative output, longer processing time, poor performance in noisy (image or sound) environments ,costing , robustness, manual update requirement and so on.

Considering these issues, the main target of this work is to develop a new method with latest technology for minimizing most of the above mentioned limitations. The complexity of real environment data association makes the object movement tracking more challenging [31]. Methods based on occupancy maps [32] produce more accurate result for still object rather than dynamic objects.

For detecting moving objects with more accuracy some systems (e.g. in [33]-[35]) extracted the object of any type by the shape modeling and consider matching of object with that. But the main limitation of these works is the requirement of enough ground visibility [37]. Category specific models using camera (e.g. in [36] - [39]), depth information [40] or both (e. g. in [41]-[43]) outputs better result comparatively. A more reliable, single camera based obstacle detection approach is proposed in [44] however simple smooth ground surface is required for this approach to wok accurately. Another obstacle detection algorithm using U-V disparity map is presented in [45] but two webcams are required for generating the disparity map. Microsoft Kinect based blind navigation system is proposed in [29] where depth image is analyzed for detecting obstacle presence and absence in the front scene and for still staircase detection only. However moving escalator would also be detected as staircase by this method which might cause dangerous accident for blind people. Again in indoor environment front scene usually contains multiple obstacles (static and dynamic). Proposed technique is an improvement and extension of existing methods like the one presented in [29].

According to natural human perceptions moving objects movement direction is more important than object or obstacle detection. When a person walk s/he concerns about moving people and other moving objects towards him/her rather than detecting and identifying what is situated around. Again an obstacle free path can change anytime for an object's arrival on that. So information about moving object's moving direction with respect to a blind person's position is more expected and important than information about all objects at the whole scene. This is appropriate for security surveillance perspective.

The proposed technique OMDIDIB (Object Moving Direction Identification using Depth Image for Blind) is a simple, affordable and realistic blind navigation support system which requires no complex algorithm or mathematical calculation to understand front scene. This system can differentiate between still staircase and moving escalators along with its moving direction (upward or downward). Any other moving object's moving direction is also detected with respect to the blind person's view point. Depth information of the font scene of a blind person is captured using Microsoft Kinect version 1 sensor. Three consecutive depth frames are taken from one second video and Distance Along Line Profile graph is generated for four lines of each depth image.

These line profile graphs are analyzed for detecting presence of any moving object and its moving direction. After detailed investigation, experimental result shows that the proposed method can successfully detect moving object along with its direction and still objects with 92% and 87% accuracy respectively. The overall accuracy of the proposed method is 90%.

4.2 METHODOLOGY

Microsoft Kinect is used for a single capture point of front scene. To detect any object movement in any direction, whole front scene image is divided into four vertical line profiles numbering 1 to 4 from right to left. Output is the information about front scene in respect to the blind person. Outcome can be delivered by any form like sound signal, vibration, pointing device movement etc. After capturing 30 frames per second configured video clip from a fixed position using Kinect, 3 consecutive depth frames are extracted, in this case, 0^{th}, 15^{th} and 30^{th} frame. Output is generated by the new OMDIDIB method. All possible outputs are listed in Table 3 and overall system overview is shown in Figure 11.

Table 3. Possible outcomes of the proposed method OMDIDIB

Front scene contains		**Type**	**Movement direction**
No obstacle / Freeway		Still	-
Wall / Door		Still	-
Staicase / Escalator	Upward	Still	-
	Downward	Still	-
Escalator		Moving	Upward
		Moving	Downward
Other Object (human, houshold accessories etc.)		Still	-
		Moving	Forward
		Moving	Backward
		Moving	Left-to-right
		Moving	Right-to-left
		Moving	Forward from left-to-right
		Moving	Forward from right-to-left
		Moving	Backward from left-to-right
		Moving	Backward from right-to-left
		Moving	In place

4.3 EXPERIMENTAL SETUP

Microsoft Kinect is widely used, available and also affordable. RGB camera and depth sensor of Microsoft Kinect is used for image acquisition for this work. Inspired by the initial setup of [29] Kinect sensor was positioned as on a standing human's chest of average heights. Ground to device height is about 1600 mm with a vertical view range of about 5000 mm (starts from 600 mm front distance of sensing device). Object distance covering range of the depth image is 800 mm to 6000 mm.

4.4 DATA COLLECTION

Since the new method presented in this chapter is capable of working in indoor environment, among many types of objects, still staircase, moving escalator, still or moving human and household accessories, wall, door and obstacle free front scenes are considered for this research work. Data are collected from open access data storage of internet and also captured with Microsoft Kinect version 1 [46, 50]. A total of 600 samples depth images of 200 different front scenes (still and moving objects in twelve different directions along with moving escalators) with their respective RGB images are analyzed to test the new system. The resolution of captured depth images and RGB images are 640×320 pixels. Resolution varies for collected images from internet which are analyzed accordingly.

4.5 PROPOSED METHOD

The method presented in this chapter is named as OMDIDIB (Object Moving Direction Identification using Depth Image for Blind) using the first letter of main words for easy referencing. Working flow of this new system is shown in Figure 12 as a self-descriptive flowchart.

Captured depth images have broken phenomenon for Kinect hardware limitations. Simple morphology processing erosion and dilation [52] is used for noise reduction. Processed depth images results better output than non-processed images as stated in [30]. After that, four vertical line profiles are extracted at pre-defined positions covering the image area.

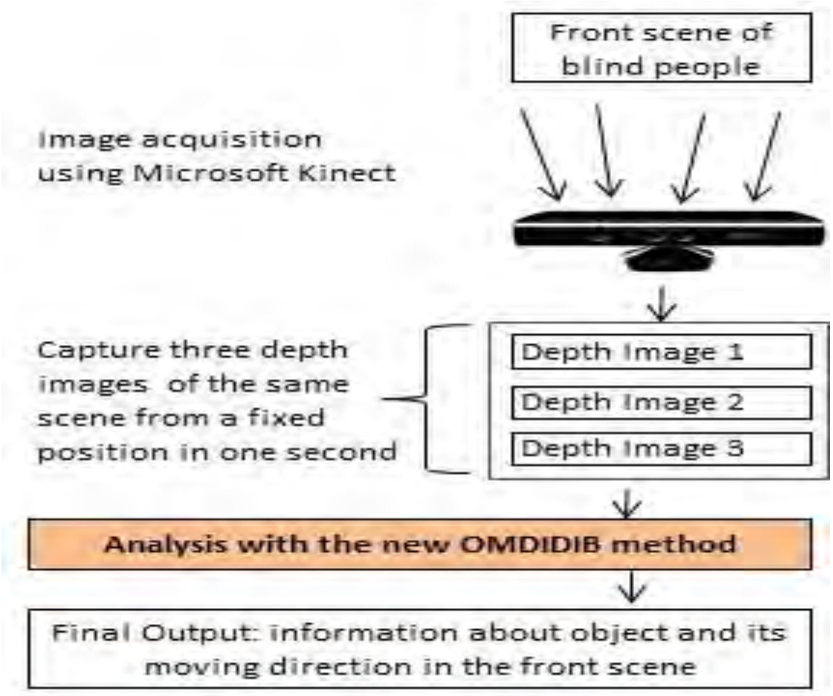


Figure 11. Overview of the overall system.

These lines are shown with number 1 to 4 from right to left in the flow diagram. Next, distance Along Line Profile graph (DAP) in short is generated in same axis range for each line profile of all three depth images. A total of twelve DAP graph is generated at this stage. Then, following six logical steps are followed sequentially.

Step 1: In all four DAP graphs of first depth image check for straight smooth upward graph shape or for sudden fall or straight up or down in graphs at regular intervals and detect upward or downward staircase or escalator from Sobel edge detected image of the original input image (gray scale image converted from RGB). Fist output is generated here deciding about any upward or downward staircase / escalator or single wall / door or no obstacle presence in the front scene.

Step 2: Detect and decide if there is any change in DAP graph for the profile lines of next images. If no change is observed output of this step is 'detected still object' otherwise the output is 'moving object is present in the front scene'.

Step 3: Similarity or matching sub graph of L length is searched in the second image under the same and different line profile graphs. L is the minimum straight length of sub-graph needs to be matched. Only the longest and best match is considered here. If matching found under the same line profile graph the movement direction is toward-backward. If matching found under different line profile graph the movement is left-right.

Step 4: Distance change of the matching sub-graph in respect to its first axis point in its previous graph is calculated here. For distance increase output decision is ‘backward’ means moving object is going far from blind person. Output is ‘toward’ if the distance increases. If match found under the graph of same positioned line profile for distance increase or decrease the moving direction is straight forward or backward. But if the match found under different DAP graph of next image, the direction is forward/backward in left-to-right or right-to-left direction. Decisions are made according to conditions shown in decision boxes in Figure 12. If DAP graph changes for any particular objects but distance remain unchanged this means the output is not still but it’s not changing its distance in respect to the user. Such output is produced for some objects like wall-clock pendulum, shaking toys, rotating ball in a single point, flying flag and so on.

Step 5: As first four graphs of first image are compared with the second four DAP graphs of second image similarly second image’s graphs are compared with the third image’s four graphs according to the above steps. More than three images can be captured for different purposes but for this work, three is selected for faster capture, analysis and response time. More capturing will produce more specific and detail output result in the price of time.

Step 6: Best outputs from different levels (five outputs at 5 levels are marked in Figure 12) of each iteration are combined and then integrated to produce the final output result. While combining the decisions first priority is given to ‘toward-backward object movement’ and second priority is given to ‘left-right movement’ as objects moving towards blind people should be notified first with highest priority. Figure 13 shows some example of input images. Detailed analysis with these images are presented below with Figure 14, 15 and 16.

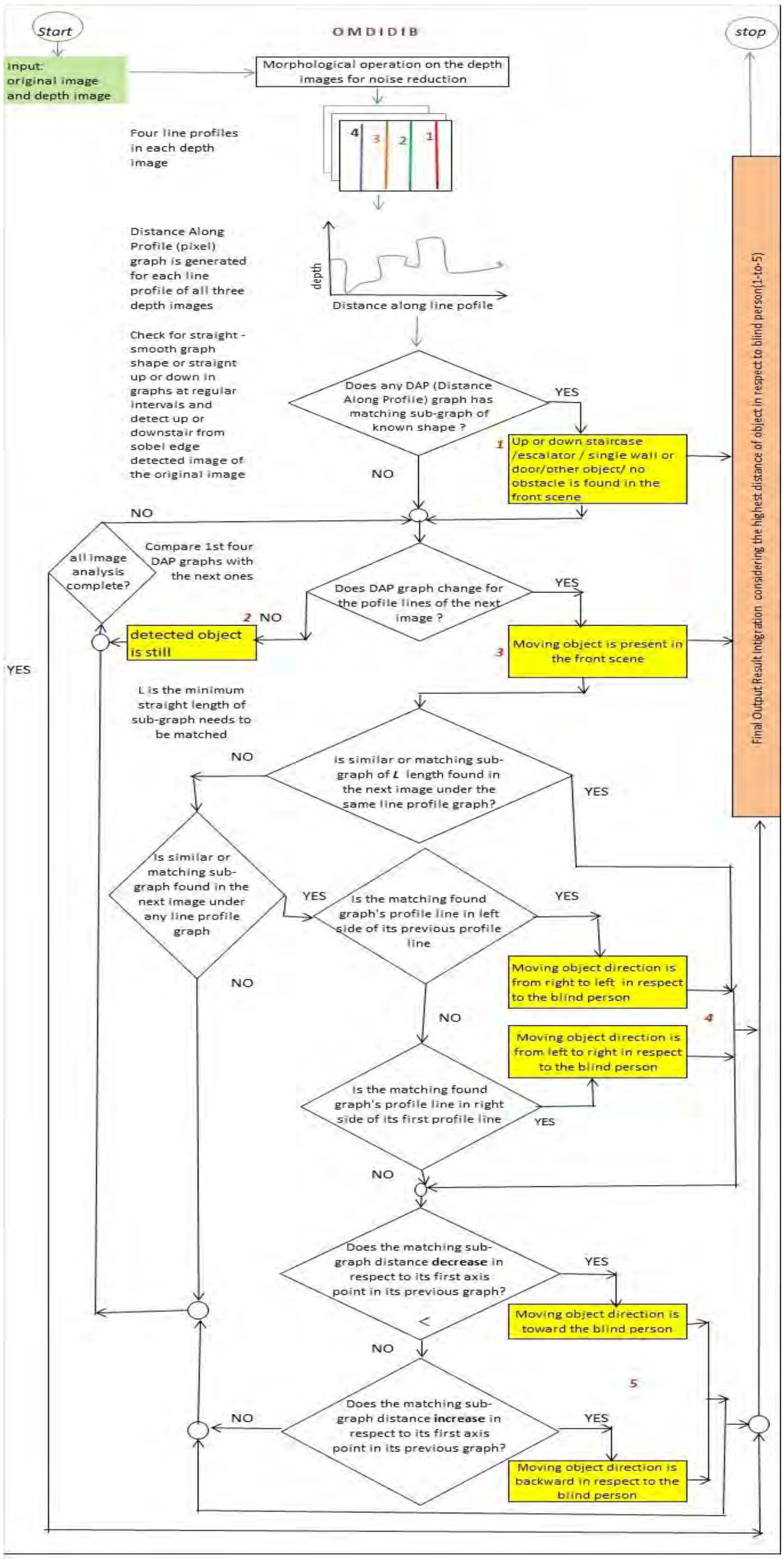
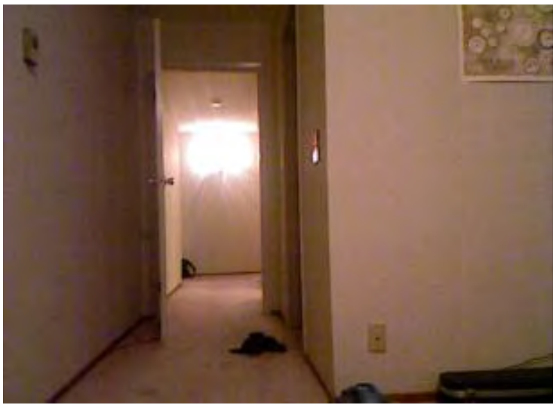
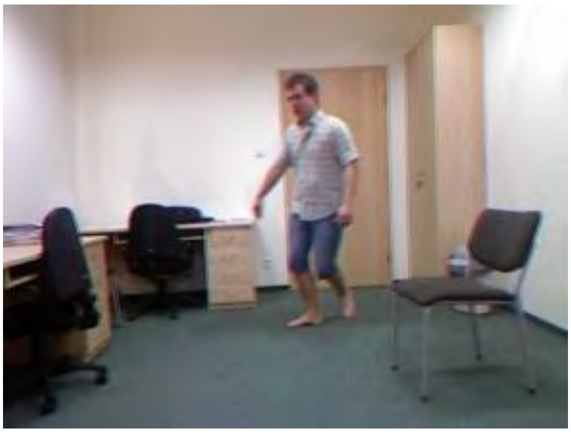


Figure 12. Flowchart of OMDIDIB (Object Moving Direction Identification using Depth Image for Blind) method.



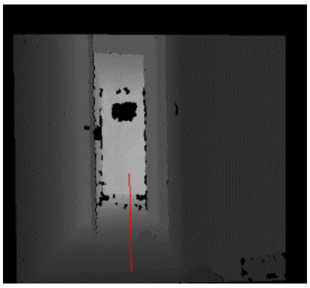
(a)

(b)

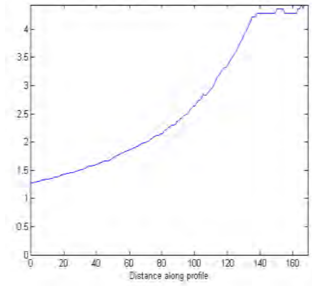


(c)

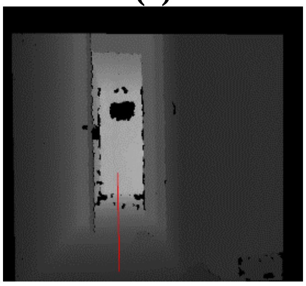
Figure 13. RGB input images of (a) freeway path along with still door [46], (b) downward moving escalator, (c) moving human [50].



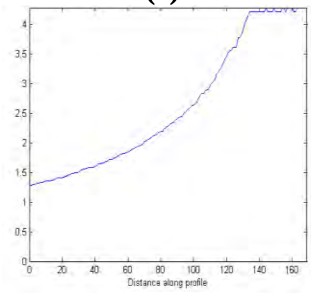
(a)



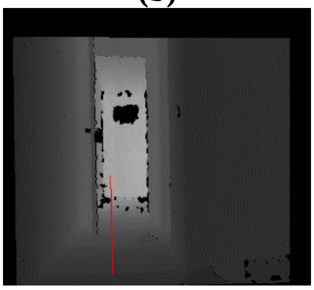
(e)



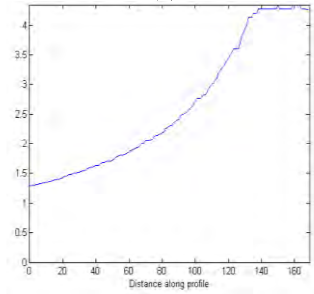
(b)



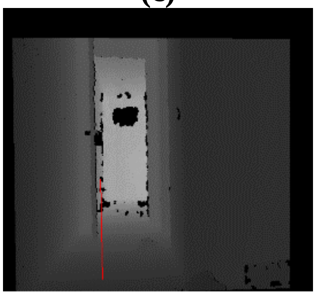
(f)



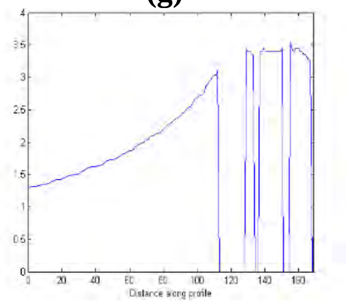
(c)



(g)



(d)



(h)

Figure 14. First one of three Depth images (a) first line profile, (b) second line profile, (c) third line profile, (d) fourth line profile of freeway path along with still door. Distance along line profile (DAP) graphs of corresponding line profiles.

Figure 14 shows one of the three depth images of freeway path extracted from video. In this single depth image four pre-defined line profiles have been extracted. Figure 14 (a)-(d) shows four line profiles separately from 1 to 4 in right to left direction, to determine the change in corresponding DAP graphs.

Figure 14 (e)-(h) shows DAP graph of corresponding line profiles. For all three depth images first three DAP graph's (e)-(g) smooth straight upward shape detects it's a freeway but last DAP graph (h) detects it is other object. Therefore combining all four line profile outputs, it detects as other object. Corresponding (e)-(h) DAP graph has no change in particular L length of other two depth images DAP graphs which indicates, it is a still object.

Figure 15 shows three depth images of moving escalator extracted from video. In these depth images four pre-defined line profiles have been extracted. Figure 15 (a)-(c) shows first line profile separately in all three depth images to determine the change in corresponding DAP graphs. Figure 15 (d)-(f) shows DAP graph of corresponding first line profile. For all three depth images DAP graph (d)-(f) straight up or down at regular intervals shape detects it's an upward staircase/escalator. Corresponding (d)-(f) DAP graph has change in particular L length of other depth images DAP graphs which indicates, it is a moving escalator. Since the change in DAP graphs are decreasing with respect to line profile, the moving escalator is coming towards or downward to blind person's view point.

Figure 16 shows three consecutive depth images of other object (a)-(c) and first depth image (d) from second scene captured video with corresponding DAP graphs (e)-(h). Similar to Figure 15 analysis, DAP graphs shows other object coming toward to blind people (e)-(g), in addition to second scene's DAP graph, L length match found in second line profile which indicates object moving from right to left direction. Combining all DAP graphs analysis, object is moving diagonal or forward right to left direction.

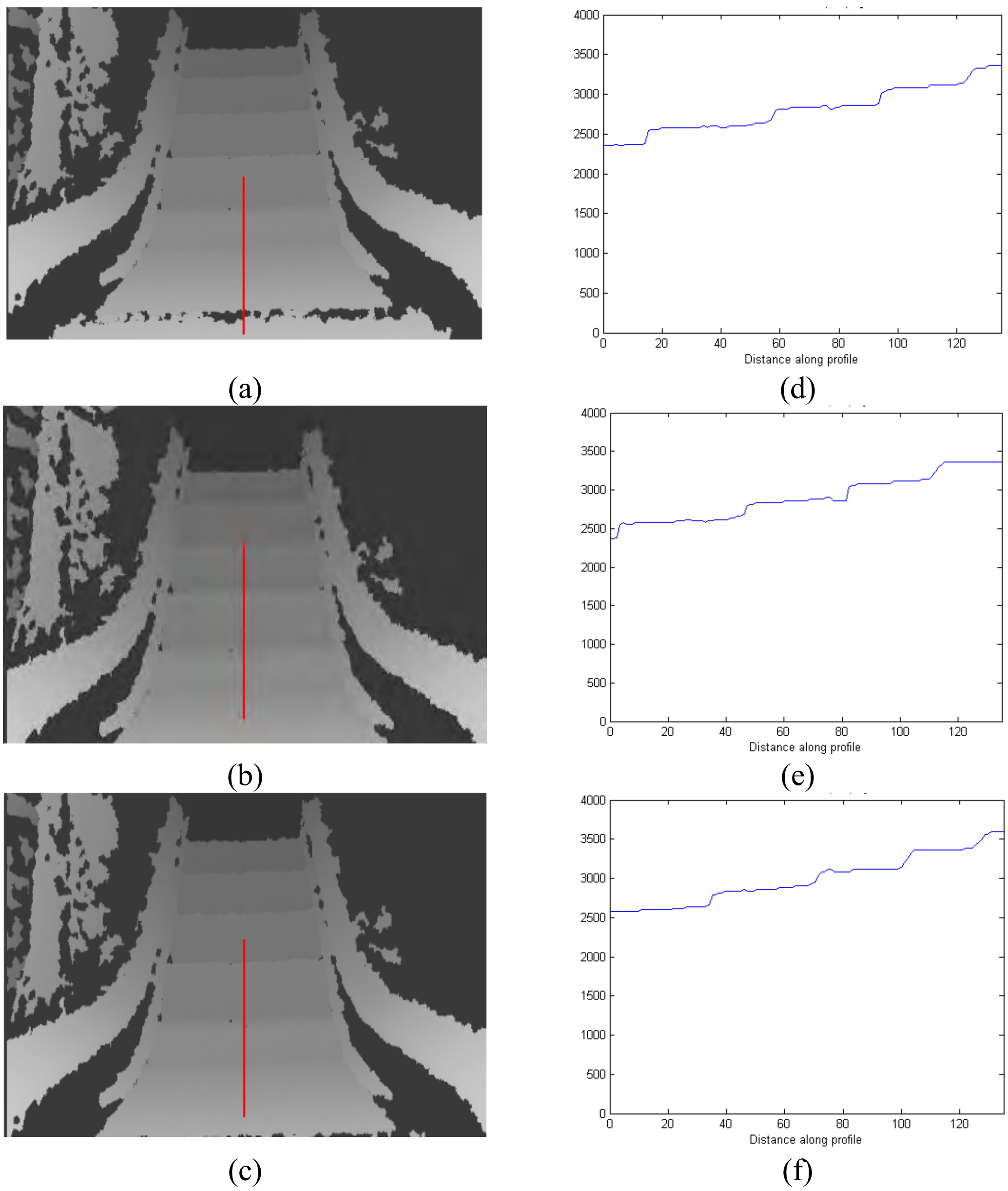


Figure 15. Single line profile of three consecutive depth images of downward moving escalator (a)-(c), Distance along line profile (DAP) graphs of respective (a)-(c) depth images (d)-(f).

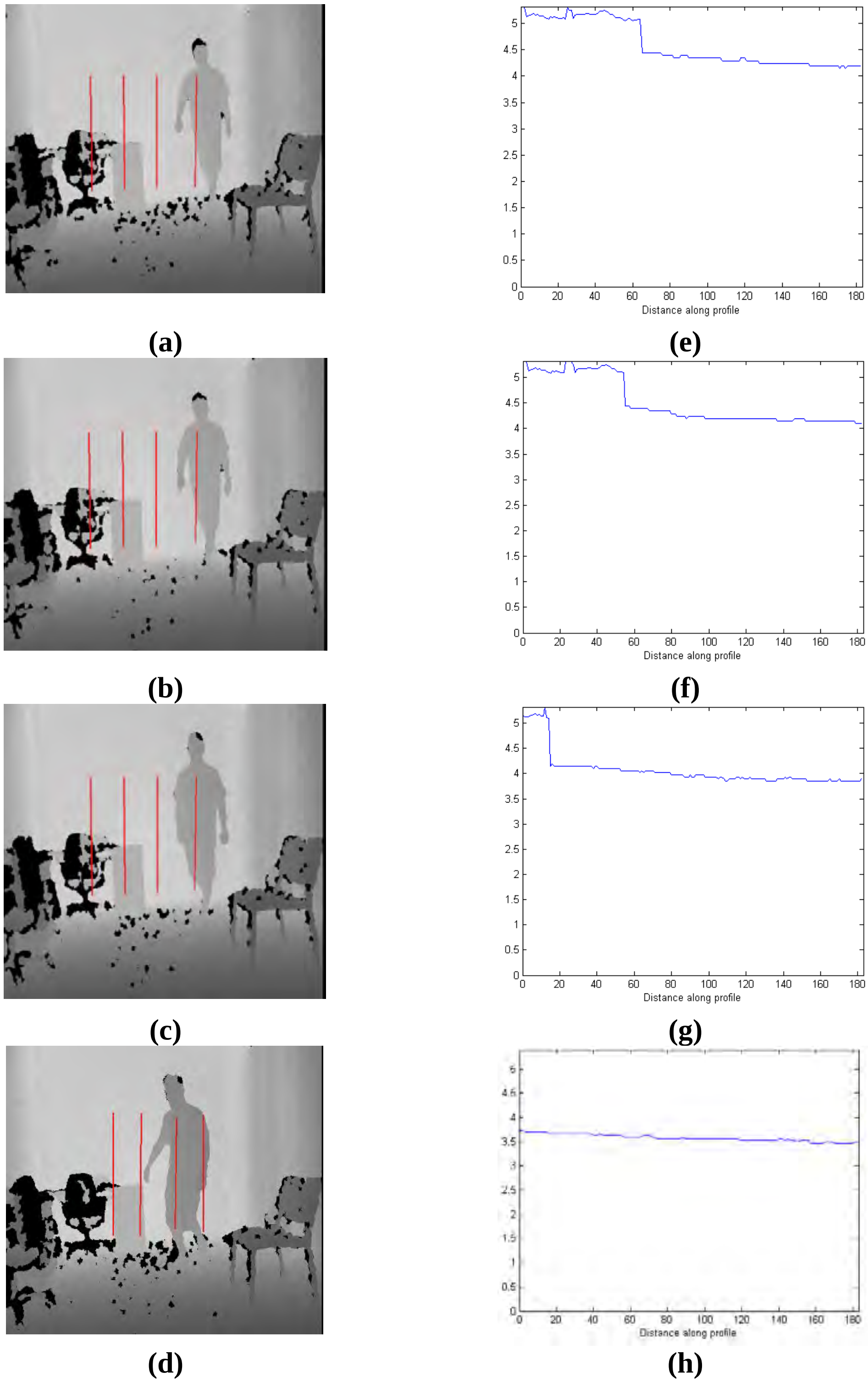


Figure 16. Three consecutive depth images of first scene (a)-(c), (d) First depth image of second scene, Distance along line profile (DAP) graphs of corresponding line profiles of depth images (e)-(h).

Table 4. Decision analysis of Figure 14, 15 and 16

Parameter	Figure 14	Figure 15	Figure 16
Object	Other objects	Staircase (upward)	Other objects
Changes in other images of same line profile	No	Yes, same changes with respect to individual line profiles	Yes, all changes are not same with respect to individual line profiles
Moving Object	No	Yes	Yes
Towards/ Backwards	No	Towards	Towards
Left/ Right	No	No	Left
Decision	Freeway ahead with obstacle on left side	Upward escalator, direction downwards	Object moving forward, right-to-left

Table 4 shows decision analysis of Figure 14, 15 and 16 respectively derived from flowchart in Figure 12. System's parameters of outputs are object detected, change in DAP, moving object, towards/backwards and left/right.

4.6 EXPERIMENTAL RESULT AND DISCUSSION

For performance evaluation of the proposed method, experimental result is investigated in two levels, obstacle identification and object moving direction identification. Success rate and failure rate of these levels analyzes are presented in Table 5, 6 and in 7 respectively. Table 5 represents the object detection success and failure rate for the proposed system. Test database is consisted of 200 different scenes taken each from one second video individually using Kinect. Three depth frames are extracted per scene, in total 600 frames are taken to analyze the system. Below tables are showing the final result of detection and movement direction per scene's by analyzing three depth frames/images. In 200 scenes, where 174 detected and identified correctly and 26 were wrong categorized which results 87% success rate. Table 6 shows the object identification success rate of 200 scenes. Table 7 shows movement direction identification of the moving objects. Table 8 shows overall system's average L length measurement with respect to line profile projection.

Table 5. Obstacle detection success rate and failure rate

	Number of sample images	**Percentage (%)**
Success	174	87
Failure	26	13
Total	200	

Table 6 shows that 87% accuracy rate is obtained for no obstacle detection where 2 sample is misclassified as wall or door and 1 as other objects due to close snapshot from Kinect. 4 samples of wall or door were wrong categorized as no obstacle and other objects with 85% accuracy due to the same reason. This may result risk for blind person thus this system requires minimum distance (614 mm) from object to identify correctly. 6 upward staircase/escalator were detected as downward staircase/escalator and other object, similarly 6 downward staircase/escalator detected as upward staircase/escalator and other objects due to light reflection and object pattern with 89% and 88% accuracy respectively. 7 other objects images were wrongly categorized due to its pattern verity with 85% accuracy. Overall system’s object detection performance accuracy is 87%.

Table 6. Obstacle identification success rate

System output	Actual output					
	No Obstacle / Freeway	***Wall / Door***	***Upward (stair / escalator)***	***Downward (stair / escalator)***	***Other Objects***	**Total**
No Obstacle / Freeway	20	2	0	0	1	23
Wall / Door	2	23	0	0	2	27
Upward (stair / escalator)	0	0	48	2	4	54
Downward (stair / escalator)	0	0	2	43	4	49
Other Objects	2	1	3	1	40	47
Rate of accuracy (%)	87%	85%	89%	88%	85%	
Average accuracy rate			87%			

Table 7. Object moving direction identification success rate

		Moving - Escalator		**Moving - Other Object**	**Total**
		Upward	***Downward***		
Moving - Escalator	***Upward***	29	0	4	33
	Downward	0	32	2	34
Moving - Other Object		2	1	35	38
Rate of accuracy (%)		88%	95%	93%	
Average (%)			92%		

Among 200 scenes, 105 scenes consist of moving objects. Moving objects are categorized as escalator (upward/downward) and other object (different directions). Table 7 depicts object moving direction identification success rate with 92% accuracy. 4 moving upward escalator and 2 downward escalator were detected as other moving objects due to ground level change.

3 other moving objects were detected as upward and downward escalator due to its pattern. Proposed method also can detect other moving direction for other objects such as, left to right, backward, forward, diagonal.

Table 8 shows the overall accuracy rate by of object identification and moving direction identification along with Length L in percentage of line profile for different objects/environments. L length is presented in percentage of line profile which differs in respect to line distance. Here L is the percentage of the line profile must match with sample images to identify the object and its direction.

Table 8. Average accuracy rate and length L in percentage of line profile for different object

Type	Object	Overall Accuracy (%)	Length L* in percentage of line profile (%)
Still	***No Obstacle / Freeway***	87	80
	Wall / Door	85	50
	Staircase / Escalator (Upward + Downward)	89	20
	Other Object	85	80
Moving	***Escalator (Upward + Downward)***	91	10
	Other Object	93	30
	Average (%)	88	45

*L is the minimum horizontal length of sub graph needs to be matched with other two images' DAP graph

This system is capable of detecting two categories of objects detection, still and moving. For detection, OMDIDIB can detect staircase and other object in separate scenes individually, more than one 'other objects' in one scene. But when staircase and object both are present in any particular scene, then staircase has the first priority to be detected. System also prefer moving objects movement timing speed to be fast to be detected as moving, otherwise L length change in respect to distance will be too short to be detected as moving, rather it will result in still object. System failure rate is 13% because of image captured partially under sunlight, Kinect cannot measure the distance. Thin object cannot be detected as it does not fall under any line profile, hence object volume size matters in some cases.

Too close snapshots can result failure. Despite of the failure rate this system can detect object and its movement direction separately 87% and 92% respectively and on average 90% which can be implemented as a navigation system and help the visually impaired people in many situations. Response time of the proposed system is faster as no complex algorithm or mathematical function is used for analyzing and generating output results.

4.7 CHAPTER SYNOPSIS

Real time object along with the moving direction for visually impaired people is presented. According to the system of OMDIDIB method, three consecutive depth frames/images are extracted using Microsoft Kinect. Four line profiles are extracted for each depth image to determine if there is any still or moving object. By using Sobel edge detection method this system can determine the captured image's pattern in four categories, staircase/escalator (upward/downward), wall/door, freeway or other objects. For category staircase/escalator, system compares each image separately to detect any movement along particular line profile. If no movement is detected along any line profiles then it outputs as still staircase/escalator otherwise moving staircase/escalator. Depending on the line profile numbers 1 to 4 from right to left it can determine the location of any moving/still obstacle in the scene. Experimental result shows that the proposed method can successfully detect moving objects along with its direction or still objects on average with 90% accuracy rate. Initial findings show promising result however further investigation is required with real life environment scenario testing with output delivering device performance.

CHAPTER V

CONCLUSION AND FUTURE WORKS

5.1 CONCLUSION

In this thesis, two different methods have been proposed to detect staircases for visually impaired people. First method is based on real time still image and second one is video taken with Microsoft Kinect. In method one, a new method is presented for up stair and down stair detection using Gabor filters and SVM. Research focus was to identify the best Gabor feature and classifier for up stair and down stair detection along with freeway and other type object. Energy features of Gabor filtered image with SVM resulted the best performance with 92% accuracy. In addition to evaluate the performance of this proposed system, four more types of features (Mean amplitude, Absolute, Real and Imaginary part) were trained individually with two more types of classifier (Naive Bayes and Back Propagation Neural Network (BPNN)). Based on the research findings, the proposed new method is implemented and tested with different test images which produced 92.9% accuracy as a system performance result on average (four categories-upstairs, downstairs, freeway and other). In method two, real time object along with the moving direction for visually impaired people is presented. According to the system of OMDIDIB method, three consecutive depth frames/images are extracted using Microsoft Kinect. Four line profiles are extracted for each depth image to determine if there is any still or moving object. By using Sobel edge detection method this system can determine the captured image's pattern in four categories, staircase/escalator (upward/downward), wall/door, freeway or other objects. After detailed investigation, experimental result shows that the proposed method can successfully detect moving objects along with its direction or still objects on average with 90% accuracy rate.

5.2 Future Works

The initial findings are promising and inspiring however, further improvements rely on increasing training dataset and real life UAT (User Acceptance Test). Limitation of method one is, it requires necessary light source to output better result. Too bright or too dark image can vary result because parallel lines of staircases cannot be detected perfectly.

Method two can be used in indoor environment only because Microsoft Kinect sensor cannot work properly in direct sunlight. However this method can produce acceptable output if outdoor environment is in a shadow zone or low day light area. Another limitation is that the output result might be affected by image illumination. These are some challenges for making this proposed system more efficient in its next version. Object distance and dimension detection along with object classification could be included with this system for further extension.

REFERENCES

[1] World Health Organization. (2014). Visual impairment and blindness (Fact sheet No. 282). [online]. http://www.who.int/mediacentre/factsheets/fs282/en/

[2] Kinect Sensor. (2011). [online]. http://msdn.microsoft.com/en-us/library/hh438998.aspx

[3] T. Kumar and K. Verma, "A Theory Based on Conversion of RGB image to Gray image," International Journal of Computer Applications, 2010, vol. 7, no. 2, pp. 7-10.

[4] M. Zöllner, S. Huber, Hans-Christian Jetter and H. Reiterer, "NAVI: a proof-of-concept of a mobile navigational aid for visually impaired based on the Microsoft Kinect," In Proceedings of the 13th IFIP TC 13 international conference on Human-computer interaction, 2011, vol. Part IV, pp. 0-5.

[5] B. Choi, C. Mericli, J. Biswas and M. Veloso, "Fast Human Detection for Indoor Mobile Robots Using Depth Images," In Proceedings of IEEE International Conference on Robotics and Automation, 2013, pp. 1108-1113.

[6] B. Enjarini and A. Graser, "Planar segmentation from depth images using gradient of depth feature," In Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems, 2012, pp. 4668-4674.

[7] C. Chen, Y. Chen, P. Lee, Y. Tsai and S. Lei, "Real-time Hand Tracking on Depth Images," In Proceedings of Visual Communications and Image Processing (VClP), IEEE, 2011, pp. 1-4.

[8] H. Nanda and K. Fujimura, "Visual tracking using depth data," In Proceedings of Computer Vision and Pattern Recognition Workshop, 2004, pp. 37.

[9] R. Munoz, X. Rong and Y. Tian, "Depth-aware Indoor Staircase Detection and Recognition for the Visually Impaired," In Proceedings of The 3rd IEEE International Workshop on Mobile Multimedia Computing (MMC 2016) in conjunction with ICME, 2016, pp. 1-6.

[10] S. Se and M. Brady, "Vision-based detection of stair-cases," In Proceedings of Asian Conference on Computer Vision, 2000, pp. 535-570.

[11] C. Zhong, Y. Zhuang and W. Wang, "Stairway detection using Gabor filter and FFPG," In Proceedings of InternationalConference of Soft Computing and Pattern Recognition, 2011, pp. 578-582.

[12] J.G. Daugman, “ Uncertainty relations for resolution in space, spatial frequency and orientation optimized by two-dimensional visual cortical filters,” Journal of the Optical Society of America A, 1985, vol. 2, pp. 1160-1169.
[13] A. P´erez-Yus, G. L´opez-Nicol´as and J. Guerrero, “Detection and modeling of staircases using a wearable depth sensor,” In ECCV Workshop, 2014, pp. 449-463.
[14] X. Lu and R. Manduchi, “Detection and localization of curb sand stairways using stereo vision,” In Proceedings of ICRA, 2000, pp. 1-7.
[15] D. Kim, K. Kim and S. Lee, “Stereo camera based virtual cane system with identifiable distance tactile feedback for the blind,” Sensors, 2014.
[16] T. Lee, T. Leung and G. Medioni, “Real-time staircase detection from a wearable stereo system,” In Proceedings of ICPR, 2012, pp. 3770-3773.
[17] Y. Cong, X. Li, L. Liu and Y. Tang, “A Stairway Detection Algorithm based on Vision for UGV Stair Climbing,” In Proceedings of IEEE International Conference on Networking, Sensing and Control, 2008, pp. 1806-1811.
[18] D. C. Hernández and K. H. Jo, “Stairway tracking based on automatic target selection using directional filters,” In Proceedings of 17th Korea-Japan Joint Workshop on Frontiers of Computer Vision (FCV), 2011, pp. 1-6.
[19] C. Morland and D. Mountain, “Design of a sonar system for visually impaired humans,” In Proceedings ofthe 14th International Conference on Auditory Display, 2008, pp. 24-27.
[20] S. Oßwald, A. Hornung and M. Bennewitz, “Improved proposals for highly accurate localization using range and vision data,” In Proceedings of IROS, 2012, pp. 1-6.
[21] Google Project Tango Developers. (2016). Project Tango Development Kit. [online]. https://store.google.com/product/project_tango_tablet_development_Kit
[22] V. Ivanehenko, J. Coughlan and H. Shen, “Detecting and Locating Crosswalks using a Camera Phone,” Computers Helping People with Special Needs Lecture Notes in Computer Science, 2008, vol. 5105, pp. 1122-1128.
[23] J. K. Kamarainen, V. Kyrki and H. Kalviainen, “Invariance properties of Gabor filter-based features-overview and applications,” In Proceedings of IEEE Transaction on Image Processing, 2006, pp. 1088-1099.

[24] L. Hong, Y. Wanand and A. Jain, "Fingerprint Image Enhancement: Algorithm and Performance Evaluation," In Proceedings of IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998, pp. 777-789.

[25] M. H. Rahman, M. R. Pickering and M.R. Frater, "Scale and Rotation Invariant Gabor Features for Texture Retrieval," In Proceedings of International Conference on Digital Image Computing Techniques and Applications (DICTA), 2011, pp. 602-607.

[26] H. Lau, K. Y. Tong and H. Zhu, "Support vector machine for classification of walking conditions using miniature kinematic sensors," Med Biol Eng Comput, 2008, vol. 46, no. 6, pp. 563-573.

[27] R. Begg and J. Kamruzzaman, "A machine learning approach for automated recognition of movement patterns using basic, kinetic and kinematic gait data," Journal of Biomechanics , 2005, vol. 38, no. 3, pp. 401-408.

[28] S. A. Bouhamed, I. K. Kallel and D. S. Masmoudi, "Stair Case Detection And Recognition Using Ultrasonic Signal," In Proceedings of 36th International Conference on Telecommunications and Signal Processing (TSP), 2013, pp. 672-676.

[29] V. Filipe, F. Fenandes, H. Fenandes, A. Sousa, H. Paredes and J. Barroso, "Blind navigation support system based on Microsoft Kinect," In Proceedings of 4th International Conference on Software Development for Enhancing Accessibility and Fighting Info-exclusion, 2012, vol. 14, pp. 94-101.

[30] Hsieh-Chang Huang, "A novel stair detection approach using Kinect sensor," Dawn OF Juanqi, 2015, vol. 1, pp. 59-70.

[31] B. Leibe, K. Schindler and L. Van Gool, "Moving obstacle detection in highly dynamic scenes," In Proceedings of IEEE International Conference on Robotics and Automation, 2009, pp. 56-63.

[32] A. Elfes, "Sonar-based real-world mapping and navigation," IEEE Journal of Robotics and Automation, 1987, pp. 49-265.

[33] S. Nedevschi, R. Danescu, D. Frentiu, T. Graf and R. Schmidt, "High accuracy stereovision approach for obstacle detection on non-planar roads," IEEE Intelligent Engineering Systems, 2004.

[34] M. Soga, T. Kato, M. Ohta and Y. Ninomiya, "Pedestrian detection with stereo vision," In Proceedings of IEEE International Conference on Data Engineering, 2005, pp. 24-29.

[35] L. Zhao and C. Thorpe, "Stereo- and neural network-based pedestrian detection," In ITS, 2000, vol. 1, no. 3, pp. 148-154.
[36] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," In CVPR, 2005, vol. 1, pp. 886-893.
[37] B. Leibe, A. Leonardis and B. Schiele, "Robust object detection with interleaved categorization and segmentation," International Journal of Computer Vision, 2008, vol. 77, no. 1, pp. 259-289.
[38] A. Shashua, Y. Gdalyahu and G. Hayun, "Pedestrian detection for driving assistance systems, Single-frame classification and system level performance," In IVS, 2004, pp. 1-6.
[39] B. Wu and R. Nevatia, "Detection and tracking of multiple, partially occluded humans by bayesian combination of edgelet part detectors," IJCV, 2007, vol. 75, no. 2, pp. 247-266.
[40] K. O. Arras, O. M. Mozos and W. Burgard, "Using boosted features for the detection of people in 2d range data," In Proceedings of IEEE International Conference on Robotics and Automation, 2007, pp. 3402-3407.
[41] A. Ess, B. Leibe and L. Van Gool, "Depth and appearance for mobile scene analysis," In Proceedings of IEEE 11th International Conference on Computer Vision, 2007, pp. 1-8.
[42] D. Gavrila and V. Philomin, "Real-time object detection for "smart" vehicles," In Proceedings of the Seventh IEEE International Conference on Computer Vision, 1999, vol. 1, pp. 87-93.
[43] L. Spinello, R. Triebel and R. Siegwart, "Multimodal people detection and tracking in crowded scenes," In Proceedings of AAAI Conference on Artificial Intelligence, 2008, vol. 3, pp. 1409-1414.
[44] Z. Yankun, C. Hong and N. Weyrich, "A single camera based rear obstacle detection system," In Proceedings of IEEE International Conference on Intelligent Vehicles Symposium, 2011, pp. 485-490.
[45] M. Zhang, P. Liu, X. Zhao, X. Zhao and Y. Zhang, "An obstacle detection algorithm based on U-V disparity map analysis," In Proceedings of IEEE International Conference on Information Theory and Information Security (ICITIS), 2010, pp. 763-766.

[46] RGB-D Dataset For Visual Odometry/SLAM Under Large Lighting Variations. [online]. http://telerobot.cs.tamu.edu/MFG/rgbd/livo/data.html

[47] Image Processing Toolbox/ Image Analysis/ Region and Image Properties/ improfile. MathWorks website. [online]. http://es.mathworks.com/help/images/ref/improfile.html

[48] Ching Wei Wang, "Real time sobel square edge detector for night vision analysis," In Proceedings of the Third international conference on Image Analysis and Recognition, 2006, vol. 4141, pp. 404-413.

[49] B. T. Shrewsbury, "Providing haptic feedback using the kinect," In Proceedings of the 13th international ACM SIGACCESS Conference on Computers and accessibility, 2011, pp. 321-322.

[50] Interdisciplinary Centre for Computational Modeling University of Rzeszow. [online]. http://fenix.univ.rzeszow.pl/mkepski/ds/uf.html

[51] Sobel Edge Detector. [online]. https://en.wikipedia.org/wiki/Sobel_operator

[52] Morphological erosion and dilation. [online]. https://www.cs.auckland.ac.nz/courses/compsci773s1c/lectures/ImageProcessing-html/topic4.htm

THE END