T.C.

ALTINBAS UNIVERSITY

Institute of Graduate Studies

Information Technologies

TRAVEL COMPANION FOR VISUALLY IMPAIRED

Abdul QADIR

Master of Science

Supervisor

Prof. Dr. Oğuz BAYAT

Istanbul, 2020

TRAVEL COMPANION FOR VISUALLY IMPAIRED

by

Abdul Qadir

Information Technologies

Submitted to the Institute of Graduate Studies

in partial fulfillment of the requirements for the degree of

Master of Science

ALTINBAŞ UNIVERSITY

2020

The thesis titled "TRAVEL COMPANION FOR VISUALLY IMPAIRED" prepared and presented by "Abdul QADIR" was accepted as a Master of Science Thesis in Information Technologies.

____

Prof. Dr. Oğuz BAYAT

Supervisor

Thesis Defense Jury Members:

Prof. Dr. Oğuz BAYAT	School of Engineering and Natural Sciences, Altinbas University	__________________
Prof. Dr. Oğuz ATA	School of Engineering and Natural Sciences, Altinbas University	__________________
Prof. Adil Deniz DURU	Faculty of Sport Sciences, Marmara University	__________________

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science.

Approval Date of Institute of Graduate Studies:
____/____/____

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Abdul Qadır

Signature

DEDICATION

I dedicate this thesis to my family and friends

ACKNOWLEDGEMENTS

I thank Almighty Allah for all His blessings and help in the completion of this project.

I would like to express my special gratitude to my supervisor, Oğuz Bayat, PhD, for all the help that he extended to me during this project. I would also like to express my gratitude towards all my teachers for their support and constant availability.

I would like to thank my parents for their constant support.

Last but not the least, I would like to thank all my friends for their support, availability and valuable company throughout my degree.

ABSTRACT

TRAVEL COMPANION FOR SIGHT IMPAIRED

Qadir, Abdul

M.Sc., Information Technologies, Altınbaş University,

Supervisor: Prof. Dr. Oğuz Bayat

Date: 2019

Pages: 52

The goal of the project is to provide such a system that can be used as a daily life tool by the Visually Impaired people to assist them in traveling safely to their desired locations and avoiding any obstacles in their path. At present, there is no low cost and simple solution which can address the traveling issues faced by the VI. “Travel Companion For the visually impaired” is a simulation-based project. TCVI takes images as input, performs object detection and recognition, maps these detected and recognized objects on the original image taken. After the processing of the images the results are displayed with the objects identified and the geo coordinates of the processed image. The VI can utilize this system to avoid obstacles in their path and get to the desired location with the help of the geo coordinates. The project successfully detects and recognizes the objects provided on the original image using surf algorithm and displays these results along with geo coordinates on the image under process. Different colors are used for the borders of the objects.

Keywords: Image processing, Visually impaired, Feature matching, Point detection, Algorithms, Object detection, Obstacles, SURF, Scenes, Computer vision.

TABLE OF CONTENTS

Pages

LIST OF FIGURES

Pages

LIST OF TABLES

Pages

LIST OF ABBREVIATIONS

TCVI	:	Travel Companion for Visually Impaired
VI	:	Visually Impaired
WHO	:	World health Organization
SIFT	:	Scale-invariant feature transform
SURF	:	Speeded-up robust features
HLR	:	Hessian Laplace region
HD	:	High definition
HT	:	Hough transform
DoG	:	Difference of Gaussian
HOG	:	Histograms of Oriented Gradients

1. INTRODUCTION

The objective of the thesis work is to provide such a solution, which can detect known objects in a certain given path. The VI (Visually Impaired) can use this solution to avoid the objects in there path. The solution will also contain a GEO location tagging which will help the user to reach their desired location.

1.1. VISUALLY IMPAIRED AND BLINDNESS

The term Visually Impaired, sometimes is also stated as vision loss or impairment , is refered to the condition of a person who does not possess the ability to see perfectly. [1]. Blindness is the condition in which a person completely losses his/her sight. [1][2]. These can not be address by routine medication or common means like glasses or lens.

The complete loss of vision is called Blindness. A blind person cannot see at all. Although Blindness is total loss of sight, sometimes it is also used as a relative term to indicate VI. For example, a person has so low vision that even with the help of medication or glasses and lenses, they still cannot see clearly.

1.2. BACKGROUND

Eye sight, clear and perfect vision is an essential part of the daily life. Due to VI The day to day tasks are effected deeply. The eye’s functions are limited and hence the everyday life of the person is accordingly effected. The clarity of vision (visual acuity) is measured by the help of Snellen’s Chart. It is a standardized chart which contains random numbers and letters. These numbers and letters vary in size. The normal vision is six by six (6/6) or 20/20. [4]

Some common causes of blindness and VI are traumatic injuries, diabetes, Cataracts, degeneration of muscles, glaucoma and cornea or retina infections. Not attending to any problems with the eyes an also lead to VI or complete Blindness. Blindness can also occur due to vascular disease, stroke, retinopathy and heredity etc. [5] As of 2015, 940 million people have been assessed to have loss of vision. [5]. The number of people with low vision was 246 million while 39 million people were completely blind. [6]. The majority of people with affected vision belong to the developing countries and are over the age of 50 years.

1.3. PROBLEMS

According to WHO, 80% cases of VI can be avoided or cured with proper treatment. [6]. The problem faced is, as mentioned earlier, the majority of the cases are from developing countries. These countries do not contain the advance technology and research to tackle with this issue. In case of the availability of surgery, rehabilitation, and the tools required for treatment, these are too expensive to afford. People with VI or blindness face problems if they enter a new environments which are not known to them. Physical challenges, like walking or traveling, in a market place or a crowded street, gets difficult. Due to this they need a companion to accompany them so that they can reach their desired destination safely. If they are walking or traveling unattended, objects in their path can seriously harm them. This may not be the case just for outdoor activites only but indoor activities may also require supervision. These People also can not contribute in their own economic growth due to lack of ability to work. It is hard for them to gain employment and is difficult to perform day to day activities.

Blindness also limits the outdoor activities like sports etc. due to which social activities of a person are also affected. They also require special software's for reading. Using the computer or a smart device can get challenging.

1.4. SOLUTION

To overcome the problem of traveling and walking, TCVI has been designed to provide ease of traveling for the Visually impaired. This system will take images as input, perform objects detection and recognition. These detected objects will be highlighted on the original image. The result will contain the original image with highlighted objects and geo coordinates of the image under process. The VI can utilize this system to avoid obstacles in their path and get to the desired location with the help of the geo coordinates.

2. IMAGE PROCESSING

"Any operations performed on an Image, taken as input, to achieve desired result is called image Processing." Or "Image processing is the study of any algorithm that takes an image as input and returns an image as output" [7].

Image is analyzed or manipulated using different mathematical operations. The input can be a single image, series of images or a video (processing is performed on frames). The result can be an image or a set of parameters of the image.

Image Processing is a computer-based technology. It can be used to provide automated solutions. It is used in numerous fields of science and technology. Following are some operations of image processing:

- Displaying image with different brightness and contrast levels.
- Image editing.
- Removing or adding objects in an image.
- Restoration and enhancement of an image.
- Feature extraction and detection.
- Image compression.
- Object detection and recognition.

Following are some applications of image processing:

- Television.
- Visual effects.
- Finger print and Face recognition software.
- Photography (Photoshop).
- Robotics.
- Industrial inspection.
- Space image processing.
- Medical image processing.

Fundamental steps, some operations and levels of image processing are shown in figure 2.1, 2.2, 2.3 and 2.4 respectively.

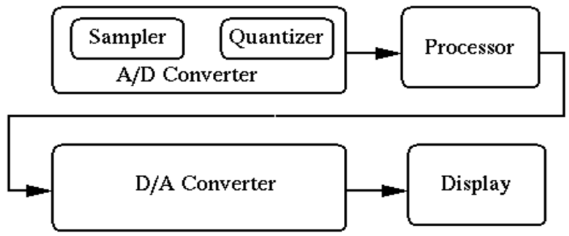

Figure 2.1: Fundamentals steps of Image Processing

Figure 2.2: Different gray levels 256, 128, 64, 32, 16, 8, 4, and 2 respectively

Figure 2.3: a) Damaged Image b) Restored Images

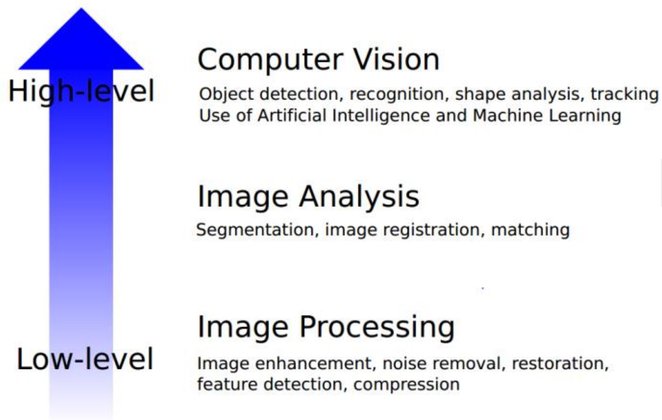

Figure 2.4: Level of Image processing

2.1. COMPUTER VISION AND OBJECT DETECTION

Computer vision is the field which deals with how a computer can be utilized to gain more high-level information and understanding from images. [8]. It provides information in the form of numerical data for better understanding. In engineering perspective, this field deals with automating different tasks which the human visual system can perform. [9].

There are numerous applications of computer vision ranging from industrial machine vision systems to robotics. From mobile applications to satellites. [10] [11]. Due to the increasing power and capabilities of both hardware and software, this field is rapidly progressing in research. [12].

In computer vision, object detection is a challenging task. It involves, taking into consideration, various light conditions, different views of the object, noise, reflection etc., which makes it hard to achieve desired results. [12]. To tackle these problems, SIFT (Scale-invariant feature transform) and SURF (Speeded-up robust features) algorithms can be used. These both feature detection algorithms are elaborated in the later section 2.2

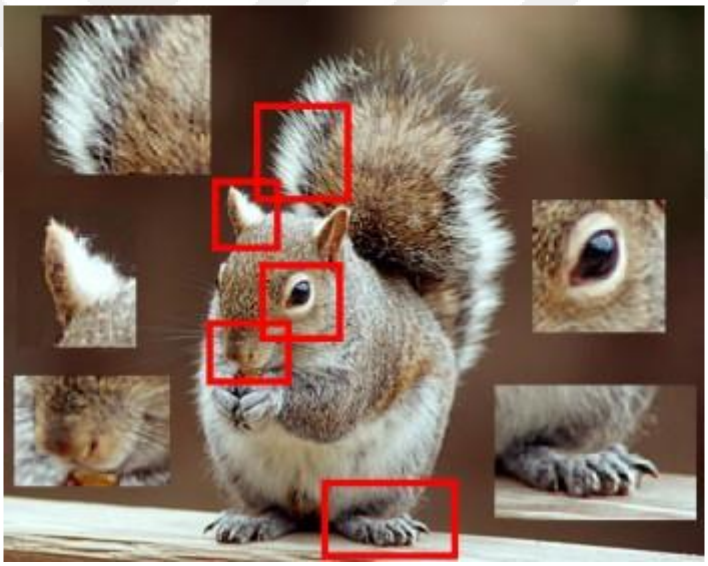

Figure 2.5: Example Image Correlation Computer Vision

2.1.1. Object Detection

The process of identifying/detecting real life objects, such as cars, faces, buildings or traffic signs etc. according to requirements, from reference images is called object detection. The algorithms for object detection commonly use feature extraction and learning algorithms to identify the objects. Some applications are facial recognition, surveillance and security. [13].

Research towards the field of feature detectors and descriptors has advanced highly due to the work on SIFT. Irrespective of partial occlusion and viewing conditions, the utilization of local invariant features to develop robust and efficient recognition provides excellent results. Commonly, image sets are used to teach the system about the objects. This set contains views of object from different angles and light intensities etc. The system uses this image set, learns from it and extracts relevant data from the image under operation.

Object detection can be summarized as following:

- Find distinctive key points.
- Outline the region of the key points at a given distance from each other. Extract and normalize the previously detected region regarding scale-invariance and orientation.
- Descriptor computation from the normalized region.
- Local descriptor matching. [14]

Invariant features are detected and selected in the first three phases. To acquire these, SURF algorithm uses HLR (Hessian Laplace region) detector along with gradient orientation-based feature descriptor to discriminate among the features selection. [15]. SIFT can use a combination of many region detectors achieving generally good performance. [16]. The analysis is continued to repeat on the image under process. The acquired features are then encoded in descriptors in the next phase. These descriptors are used for the oncoming matching, where two feature sets The same analysis is then repeated on the analyzed image, or video frame. In the next phase the obtained features are encoded in descriptors; such descriptors are then used in the upcoming matching where the two feature sets are dealt with to acquire the similarities. If it is true according, with respect to a defined threshold, object is detected. [16].

Some examples and techniques are shown in figure 2.6, 2.7, 2.8 and 2.9.

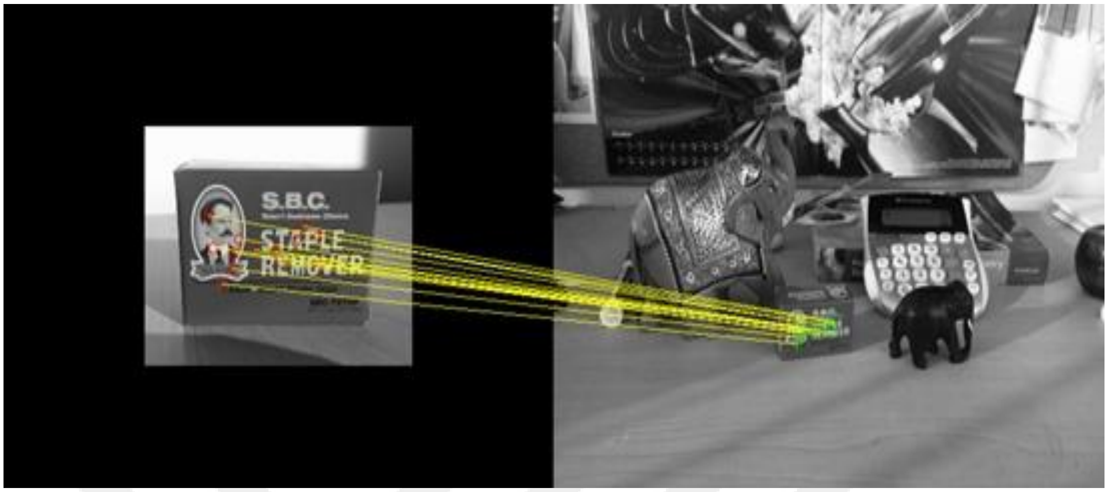

Figure 2.6: Feature Based Object Detection

Figure 2.7: Voila-Jones object Detection

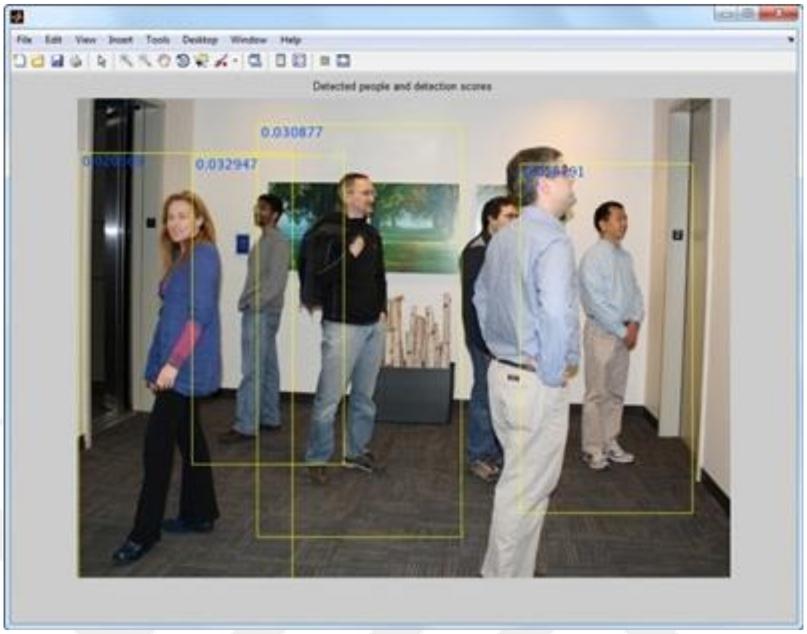

Figure 2.8: SVM classification with HOG features

Figure 2.8, using SVM classification the humans are detected with different labels and identity. For the accurate results the objects data must be trained from different angles.

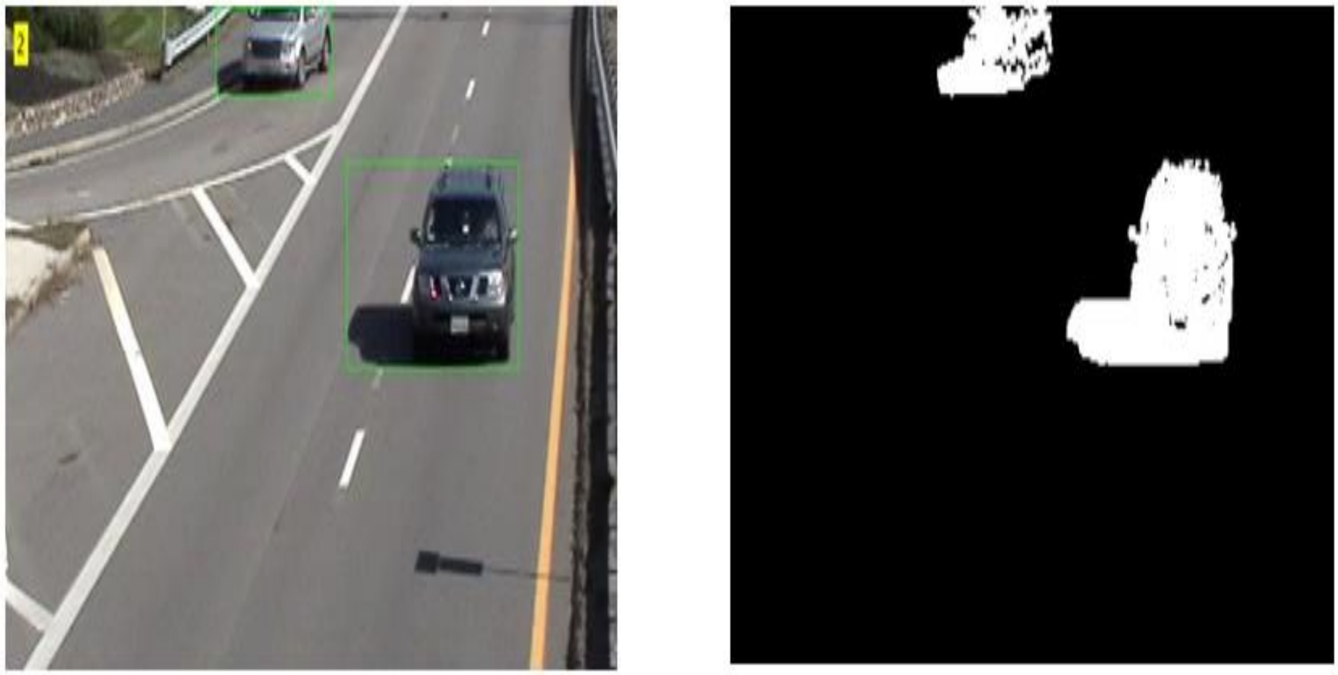

Figure 2.9: Image segmentation and Blob analysis

The Blob Algorithm is applied for the separation of objects in the scene, in cases where one scene has many different objects in a single picture for this purpose the objects separation is mandatory. In figure 2.9 the Blob algorithm is used for the segmentation purpose, in figure 2.9, image is separated in two parts, and then the result image will be processed for matching purpose.

3. LITERATURE REVIEW

The literature review is focused on different methods of developing a system that can help VI in their basic need of traveling. The literature covered includes the current products available, research papers concerning object detection and recognition which will assist in the by the project group included some medical books and research papers that will help in the achievement of the task at hand. The abstracts of some of the research papers are stated below.

Tarun Agarwal, [17] discusses the current utilization of robot's due to their reliability and high-level performance. This results in great advantage for human beings. The aim is to use a robot to detect objects/obstacles and avoid any contact or collision from them. The robot will perform this operation autonomously. To make the robot capable to perform this task, variety of sensors are required. The sensors are integrated according to their tasks. The key task in this paper is to achieve object detection. The robot acquires the essential information from the surrounding environments via the integrated sensors. Some sensors used are bump sensor, infrared sensor, ultrasonic sensor etc.

RLabayrade, D. Aubert and J. Tarel, [18] proposes a solution to provide a fast and robust method to detect objects without using a flat surface geometry assumption. In general, roads and streets are not flat. Hills and valleys can be present. They roads can we steep or rising. This results in numerous errors in precision of detecting the true object and its location. These problems can also lead to false object detection as well. Most of, many of the techniques available, for object detection, consider the road to be flat. The method proposed in this paper will be able to detect objects irrespective of downhill, uphill roads. The approach for achieving this is comprised on the construction and investigation of the v-disparity image that will provide better representation of the geometric content of the street or roads. This image will provide favorable semi-global matching which will result in robust object detection, even in the case of any partial occlusion or committed errors during matching, which will include the relative height and pitch of the stereo sensor with respect to the road surface. The objects present above the road surface are located according to the estimation of their longitudinal profile and are extracted as obstacles resulting in the accurate objects detection.

In this project, [19] the author provides a solution which is comprised of both hardware and software integration. The system uses ultrasonic sensors, with the help of which, it transmits signal which covers an angle of 180 degrees. After colliding with an obstacle the signal comes back to the system. After this the certain area from which the obstacle was detected that area in the (Blind Man Companion/Smart Shoe) will provide a vibration and a beep notifying the bearer to take a certain action i.e. move away from that point.

Tam Nguyen, Josh Slonaker and Mohammed Kadous, [20] proposes a solution for real time object detection in a dynamic environment. The motivation for this paper is the increasing interest in computer vision and its applications. The system, in its final stage, will be fast and robust, which will allow the robot to adapt according to any environment change, with nominal time. To acquire HD (high definition) input, the system uses Microsoft Kinect. State of the art algorithms like SIFT and SURF are used for analysis of object recognition and machine vision. The suitable algorithms have been optimized by continuous testing for the desire hardware.

Duy-Nguyen T, Wei-Chao Chen, Natasha Gelfand and Kari Pulli, [21] describes the importance of robots in the modern society. The applications of robots are in a variety of fields. They range from medical applications to industrial. From military to manufacturing. The aim of this project is to add a degree of autonomy to an already functioning user-controlled robot. The robot designed is comprised of ultra-sonic sensor proximity range sensor, which is carefully interfaced to get the full functionality accurately, so that it can obtain input from the surroundings and perform its task without any obstruction. This sensor is used for object detection in the path of the robot. The robot is also equipped with GPS device to get the current position of the robot and navigate it to the desired location.

F. Ren, J. Huang, R. Jiang, R. Klette, [22] proposes an algorithm which can efficiently recognize image continuously and track feature descriptor of a given input video. This algorithm functions by reducing the search space of possible interest points inside of the scale space image pyramid. It performs searching and matching of candidate features in local neighborhoods, inside the 3D image pyramid. It does not compute their feature descriptors. These candidates are validated further through a motion model. Descriptor computation becomes efficient because only those areas of the image pyramid that contains the features are searched. The final state is demonstrated on real-time object recognition and label augmentation running on a mobile device.

In this article, [23] proposes a method to recognize general traffic signs. The method uses SIFT/SURF algorithm for feature matching. The colored input images/frames are first converted from RGB to HSV color space. Afterwards different shapes like circles, triangles or squares, which can be potential sign, are detected using HT (Hough transform). After this, these signs are compared with predefined reference signs from the pre-made database using SIFT/SURF algorithm for feature matching. The traffic signs are recognized as the result. After experimentations, the proposed method was robust for the data under process with 95% success rate.

3.1. CONCLUSION

The conclusion made after reviewing these papers and projects clearly defines the ability to provide new system or technology for VI, which can be used as an alternate option. After going through literature and doing research, the aim of the thesis was to provide such a system that can be used as a daily life tool by the VI people to assist them in traveling safely to their desired locations and avoiding any obstacles in their path. This system will take images as input, performs object detection and recognition, maps these detected and recognized objects on the original image taken. After the processing of the images the results will be displayed with the objects identified and the geo coordinates of the processed image. The VI will be able to use this system to avoid obstacles in their path and get to the desired location with the help of the geo coordinates.

4. PROPOSED TECHNIQUE FOR THE SELECTION OF SOLUTION

There are many different methods available for solving the task at hand. To fulfil the requirements of the project following are the points which were needed to be satisfied.

- Robust: The quality of the feature detection should not be compromised irrespective of scaling, rotation and noise etc.
- Accuracy: The features should be localized accurately.
- Recurrence: The algorithm should be able to detect the same features of the same image under process, irrespective of different viewing angles.
- Efficiency: The algorithm should be efficient.
- Quantity: Maximum feature should be detected.

After Research and discussion, to ensure the simulation will meet the requirements, SURF algorithm was selected. MATLAB was selected as the tool to perform the development.

4.1. MATLAB

Being a powerful tool for image processing, all the operations and image processing was done in MATLAB. Its image processing tool box contains a vast set algorithm for image processing. MATLAB contains high-level scripting and supports variety of different libraries which makes development fast and easy. Its memory management, code simplicity and debugging fulfills the project requirements. [24] [25].

4.2. SURF

SURF is the fast approximation of SIFT (Scale-invariant feature transform). It is a fast and robust algorithm for feature detection. It consists of three parts.

- Interest Point detection:
- Local neighborhood description.
- Matching. [26].

Figure 4..1 shows the flow of SURF algorithm.

Figure 4.1: Flow of SURF algorithm

Surf locates the interest points. These are the points where maximum variance occurs. This step is usually achieved by comparing DoG (difference of Gaussian) in each location of image under different scales. After this, it constructs a vector around the previously found features for extraction. Afterwards for matching SURF follows the following steps:

- Take a descriptor from image under process.
- Compare this with object images.
- Locate the nearest descriptor with the minimum distance with the image under process descriptor.
- Store the nearest descriptor distance.
- Repeat step two and three with a new descriptor.
- For average distance, add the distance of the stored descriptors and divide by the total descriptors of the image under process descriptors.
- Store the average distance
- Go to step one for every object image.
- Compare the stored average distance. The minimum will be the best match with the Image under process.

Figure 4.2 and 4.3 shows same objects detected from the same scene with a different view angle.

Figure 4.2: Object detection using SURF view 1

Figure 4.3: Object detection using SURF view 2

Figure 4.4 shows object detection using SURF in which the critical points in each super pixel are estimated using the Accelerated Robust Feature (SURF). These key points are then used to perform matching tasks for each detected key point of the scene in the estimated super pixel.

Figure 4.4: Object detection using SURF view 3

In computer vision SURF is patented as local feature detector and descriptor. Figure 4.4 used SURF detector which has been used to detect and analyze objects. In figure 3 the interest point is detected first, which is the sign board using SURF algorithm for more better result the Blob detector is used for the same image for approximate values. The feature descriptor of the objects is based on sum of the response around the point of interest. SURF algorithm is categorized in three parts as point of interest detection, local neighbor description and matching.

5. PHASES, EXPERIENCES AND PROBLEMS

5.1. PROJECT BLOCK DIAGRAM

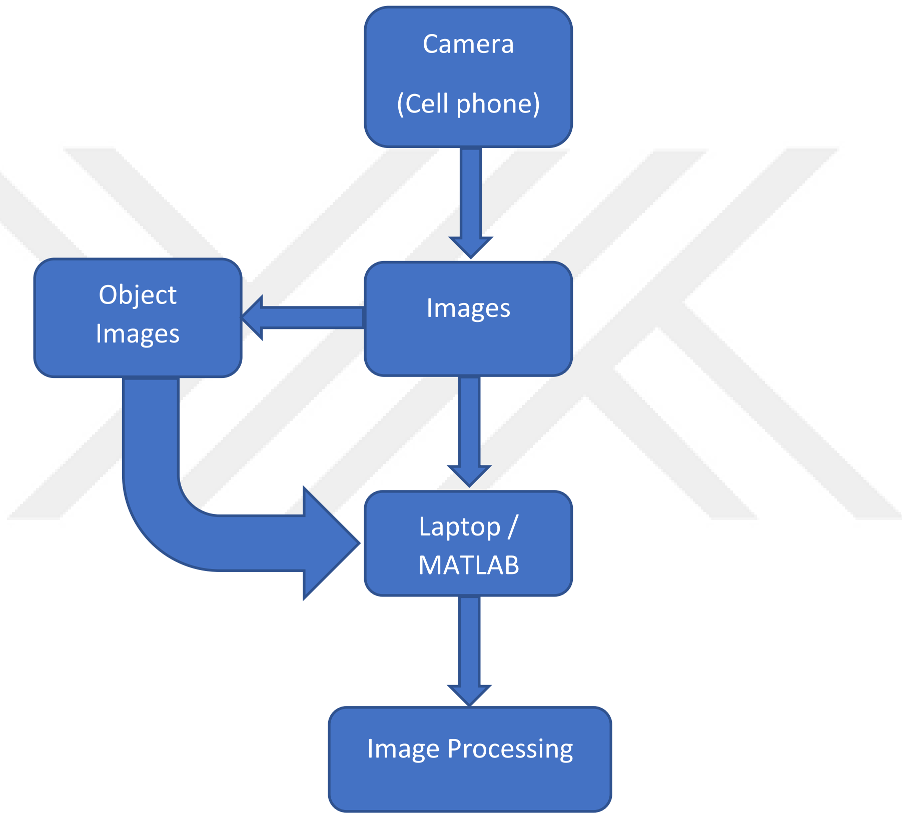

Figure 5.1: TCVI Block Diagram

5.2. PHASES

The project was completed in phases. The project was divided into the following phases.

- Basic Concept
- Platform Selection
- Research
- Data Acquisition
- Image Processing
- GEO Tagging
- Results Presentation.

Figure 5.2 shows the block diagram of the phases.

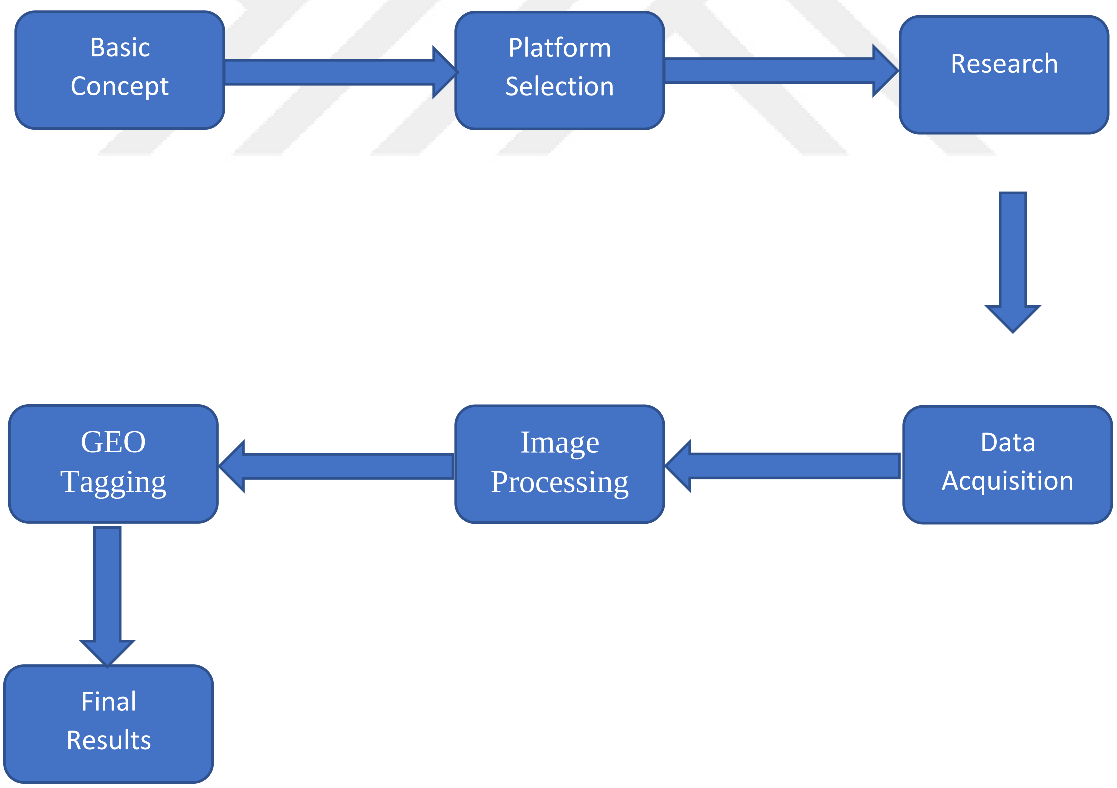

Figure 5.2: Block Diagram of Phases

5.2.1. Basic concept

The basic concept of the project is to design such a system that can assist the VI to travel. The system would detect objects, so they can be avoided, a path which can be used to travel and assign GEO coordinates, so that the user can arrive at the desired location. For this purpose a simulation was designed to fulfill the proof of concept. Images (scenes) were taken as input for the system. Different objects were selected to be detected from these scene images. After detection these objects were bordered with colors and GEO tags were shown on the final image.

5.2.2. Platform and Environment Selection

The platform selected to complete the project was MATLAB. It is explained in section **4.1 MATLAB**. The environment chosen was Heidelberg. All the scenes used in the project were capture from this street. All the objects detected were taken from these captured scenes.

5.2.3. Research

The research involved for the completion of this project contained different projects, products and research work which could addressed the task at hand. Some papers and projects are discussed in section **2 Literature review.**

5.2.4. Data Acquisition

The initial scene images were acquired using google maps. After successfully detecting and mapping objects on the final result image, cell phone camera was used to acquire further scene images from Heidelberg. From these scene images different objects were selected for detection. These both were given to the system as input for processing. For this project, forty eight (50) objects were detected from twenty eight (30) different scenes.

5.2.5. Image Processing

Different operations were performed on images, taken as input, to reach the final results. This is simply referred as Image processing. Its applications are vast and are used in many fields.

After research and discussion, the algorithm selected for the completion of this project was SURF. SURF is explained earlier in section **4.2 SURF**.

5.2.6. GEO Tagging

After successful object detection and their mapping, the final step is to add, on the final result image, is GEO coordinates. These GEO coordinates were taken from the scene image properties, when they were taken. The coordinates were extracted and the result images were tagged accordingly.

5.2.7. Final Results

The final result contains all the scene images, in which the detected objects were bordered with different colors (for our understanding), green being a safe path and red being object to be avoided. These final result scene images also contained GEO coordinates.

5.3. EXPERIENCE

The experience of the project was educating. There was a lot of exposure towards computer vision and image processing. Not only did i learn new techniques but the method of researching and completing different task on time and in a disciplined was also achieved. Due to this project, and working with my peers and teachers, it gave me an overall confidence. I was introduced to new platforms and technologies, which gave me the satisfaction of making the right decision to choose this university.

5.4. PROBLEMS

There were some problems faced during data acquisition. A lot of parameters were taken into consideration during this i.e. clear weather and sunlight etc. Besides this, from start to finish, the whole process went smoothly.

6. METHODOLGY / EXPLANATION OF THE REALIZATION

6.1. IMAGES

For the initial testing, images were gathered from google maps. Figure 16 shows the initial scene image. For better result first the trained objects taken from different scene are matched with the google images at different location where these objects exist.

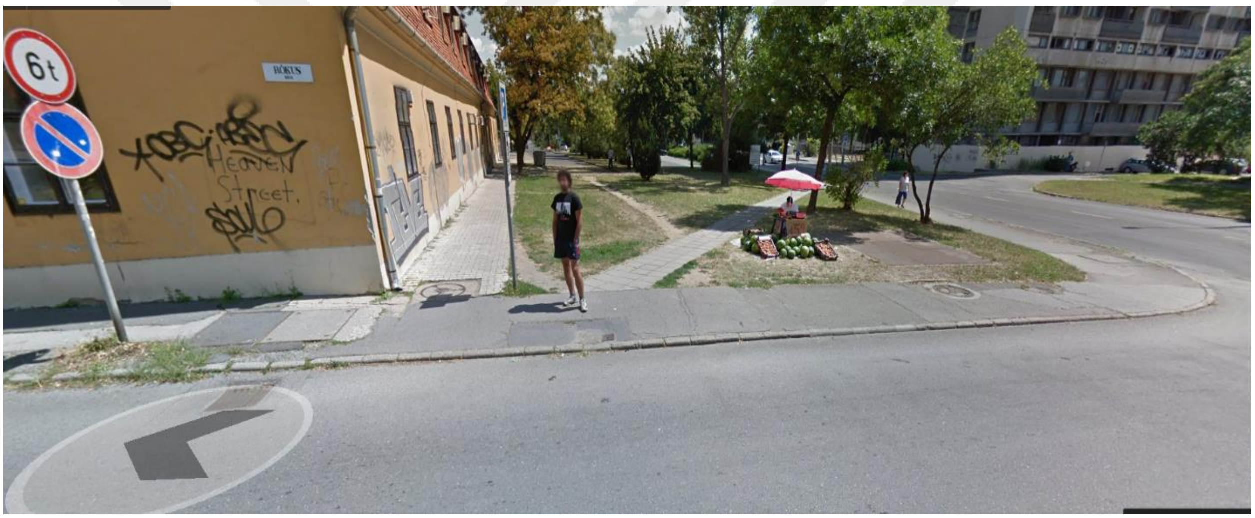

Figure 6.1: Initial Scene for testing

Different images were taken, from different viewing angle of the same environment. The objects, for detection, taken from the scene image are shown in figure 17 and 18.

Figure 6.2: Objects for detection from initial Scene (A)

Figure 17 shows the trained data objects for matching with scene in figure 16, for the better objects detection from scene pictures are taken in a close distance with clear view of objects. The objects human, pole and sign board would be detected depending on the provided data from scene as shown in figure 16.

After successful object detection from the initial scene, pictures were taken by a smart phone. Few scenes were taken after every few meters in Heidelberg. The objects for detection were taken from these scenes. The scene's used in the project and there order are shown in figure 18. The objects from the scene are detected accordingly from the train data and for multiple objects detection from one scene blob analyse and segmentation is used. Figure 18 below contatins one objects of interest in a scene in which we don't need segmentation these kind of scene are more easier to analyse as compare to multiple objects images.

Figure 6.3: Scene Image with single object

In figure 18 there are four scenes each containing one object with clear view. The targeted object detection from above scene are a car , no parking sign and map before processing these object firsts we need to train them separatly. Training of scene data from single objects are more easie as compared to multiple objects and take less time for extracting the key points.

Figure 6.4: Matching objects for Scene Images

Figure 19 shows the single object images for detecting the targeting match from scene, these are the input images for matching with the original scene, there are four scenes in figure 19, each containing one object, in the first part the map object will be sort out from the scene in figure 18, in the second part of image the program will seach for the car object in the scene. The targeted object detection from above scene are a car , no parking sign and map before processing these object detection first we need to train them separatly. In figure 20 each scene includes different objects for this purpose before analaysing the image , the objects from the scene needs to be separte using segmentation and then compare these segmantized objects with the original image.

Figure 6.5: Scene images with multiple objects

In figure 6.5 the images are shown with multiple objects in most cases, there is a single scene with multiple objects, in these types of images it’s difficult to detect all the objects. In order to find all the possible objects in figure 6.5, more data needs to be trained for matching purpose with the original scene.The scene which contains multiple objects, in the first part the map object will be sort out from the scene in figure 6.5, in the second part of image the program will seach for the remaining matching. In figure 6.6 below there are different scenes with irregular shapes.

Figure 6.6: Scene Images with irregular shapes

Irregular shapes are more difficult to detect because it’s very difficult to differential the similar objects in one scene. Figure 6.6 shows the Scene images with irregular shapes in which there are different walls anb object with same attributes, to our come this problem the blob algorithm is used which divide the scene objects into differetn segments after segmentation each object is detected individualy. The objects from the scene with different side of image provides more accurate results. In figure 6.7 below the input screen contains additional information that is not required for classification. For this reason, the image is first simplified by the scene by subtracting the important information from the image and the rest is missing. For example in figure 6.7 , if we search for the flag and logo in the image, you will notice a significant change in the RGB pixel values. However, we can simplify the image by running the edge sensor in the image and remove the information which we don't need.

Figure 6.7: Scene images with regular Objects

Figure 6.7 shows all the scenes which contains clear objects with normal shapes. These objects are very important as normally these types of obstacles are everywhere during the travel. To find these types of objects first we need to find real-world object instances (such as cars, bicycles, sign boards, poles, and people) in still Scenes or videos. It recognizes, locates and detects multiple objects in an image, which allows us to better understand the entire Scene. Figure 6.8 shows the remaining objects from the scene.

Figure 6.8: Scene images

The objects, taken from scene images are shown in figure 6.9. All these objects are inputs for the program which are matched with the scene images. The scene and object both images key points are extracted using surf algorithm and then matched for further process.

Figure 6.9: Object images for classification

Mosty the objects shown in figure 6.10 are commonly used in many places which is very important for Visually impaired people to find the path and places. For these type of objects the data of images is trained from different side in order to get more accurate results. Before the objects trained the input image is preprocessed to normalize contrast and brightness effects. A very common pre-processing step is to center the image density and divide by the standard deviation. Sometimes, gamma correction produces slightly better results. Color space conversion can help you get better results when working with color images.

Figure 6.10: Object images from different angles

For images with complicated and similar structures it's very challenging to reongnize the objects from one angle, in this case the objects are trained from different sides before matched to real world scene images. Figure 6.10 shows different object images from different sides in order to detect them accuratly. Figure 6.11 below shows the regular objects which contains only important information or main key points , these object feature are did not contain irrelevent information and need no pre processing for training.

Figure 6.11: Regular Objects with key points

6.2. MATLAB

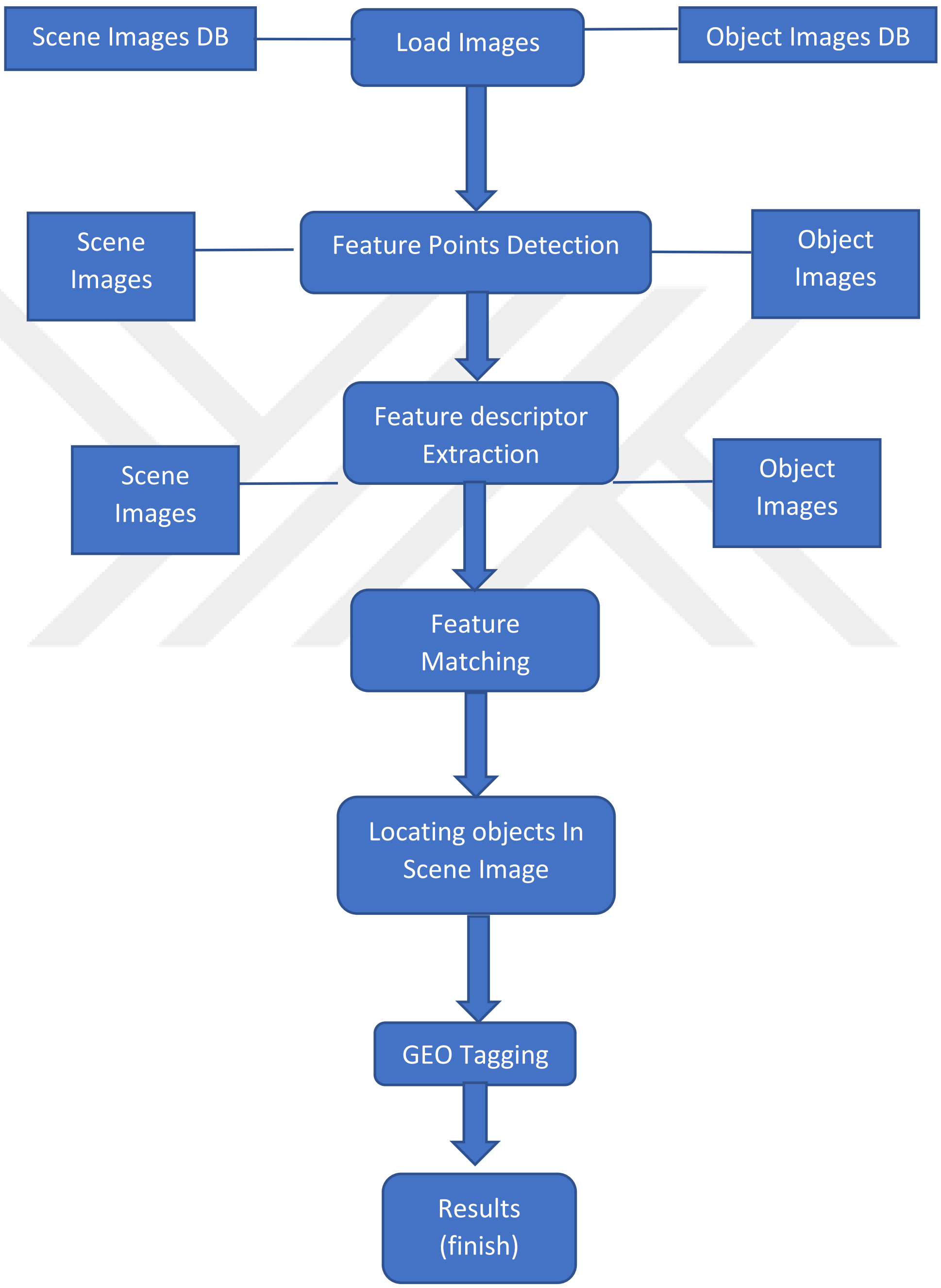

Figure 6.12: Block diagram MATLAB

6.2.1. Load Images

For the project, pictures were taken using cell phone camera. These pictures covered different scenes which contains objects that can be useful for Sight Impaired. The idea is to detect objects from these scenes. These pictures were given to MATLAB as input along with the object images. The objects to be detected and locate on the scene images were taken from these scenes. These objects images were dealt as object image database and the scene images were dealt as scene image database. The scene and object images are shown in figure 6.11 and figure 6.12 respectively in section 6.2 Images.

All these images, after loading, were converted to grayscale for further operation.

6.2.2. Feature Points detection

After loading and conversion to grayscale of scene and object images, feature points were detected from both images using SURF method. The method was implemented as a function which returned a SURF key points object for every image containing matching feature information. After this, for visualization purpose, 200 strong feature points are displayed, in figure 6.13, 6.14 and 6.15, of both the object image and scene image.

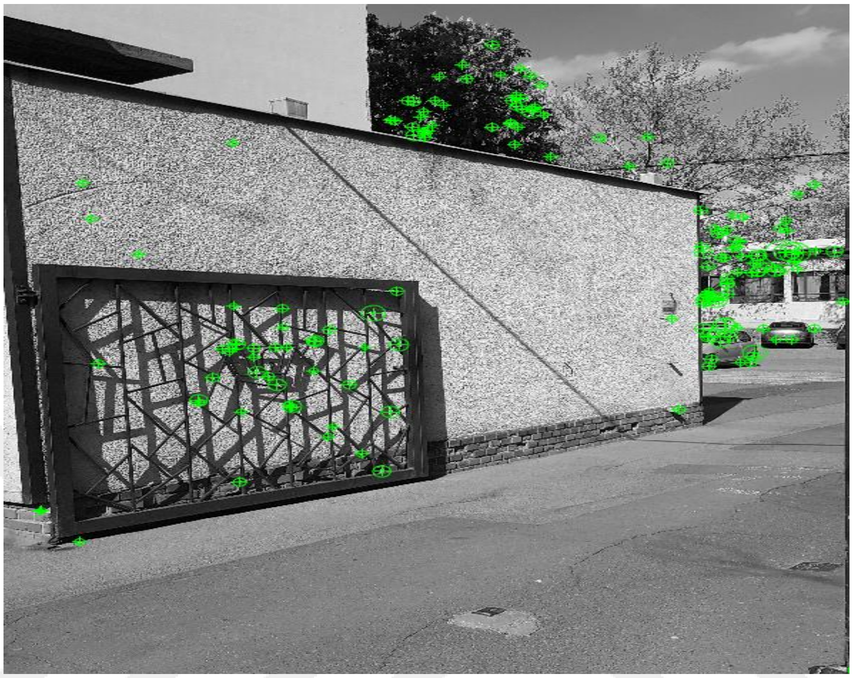

Figure 6.13: 200 feature Points Scene Image

Figure 6.14: 200 Feature Points Scene Image

6.2.3. Feature descriptor Extraction

The next step is to extract feature descriptors. This function of the program took in SURF points object of both the scene and object images as input and from these returned feature descriptors. These were further sent to feature Matching function. Features descriptor indicates the location of the important field in the image area and image. An example is a corner detector that removes the position of a corner in the image but does not say anything about the detected function.

6.2.4. Feature Matching

In this step, previous results were used to match the common features between the different objects and the scene image. Feature Matching is a simple picture mode in which we can define what can be seen on the picture. In Figure 6.15 the gate object from the scene is the feature in the image. The fundamental role of image processing is the features transforms and get visual information into vector space. This information is used later to perform mathematical operations. Feature matching is shown in Figure 6.16.

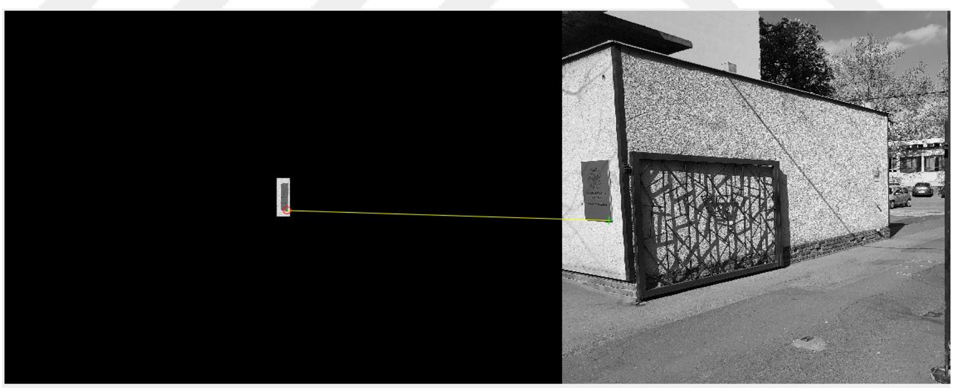

Figure 6.15: 200 Feature Matching in Two Scenes

In figure 31, the attribute mapping algorithm first determines the key points in the image which are normally the center points of the image. For extracting key points SURF algorithm is used for both images. The number of key points may vary from image to image, the greater the key points matching will provide more better accuracy. Therefore, we add 200 Feature Matching so that the attribute vectors always have the same size. Then we create a vector identifier based on our key points.

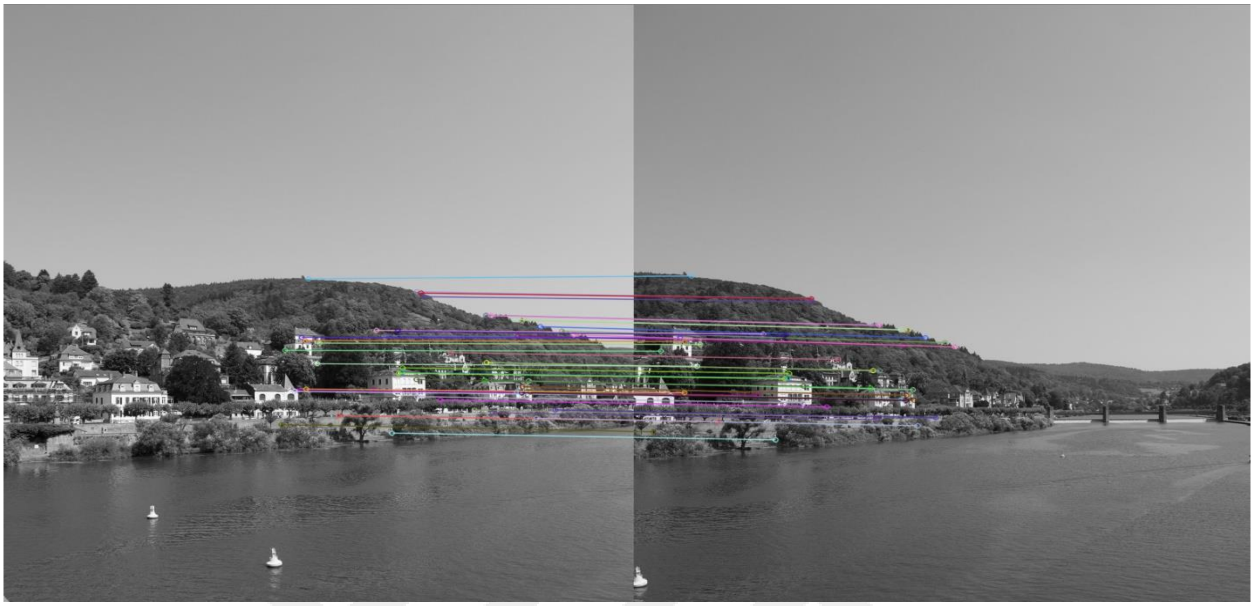

Figure 6.16: key Points Matching Between Two Scenes

Figure 6.17: Object key Points Matching with Scene key points

6.2.5. Locating objects In Scene Image

Finally the detected objects are mapped on the scene image in this step. A bounding polygon is generated for each object image according to their size. These are then transformed according to the scene image coordinates. These transformed polygons specifies the object location in the scene. Detected objects are shown in figure 6.17.

6.2.6. GEO tagging

At this point all the objects of interest are detected. The final step is to provide the scene, containing the detected objects and the GEO coordinates required. The GEO coordinates for every scene image were taken from the scene image properties. The coordinates were extracted and added to the respective scenes accordingly. Figure 6.18 shows the detected objects and GEO tagging on the scene image.

Figure 6.18*:* Geo Tag Scene Image

7. RESULTS

We have tested and trained different objects from the scene at a time by applying the Blog algorithm, SVM, Feature Matching and SURF. The TCVI measures an obstacle from normal distance by using matching techniques such as Feature descriptor and Feature methods. Then the SURF algorithm observed the objects in the scenes on different surfaces. In table 1 the detected objects data are recorded with detection accuracy and failure [27]. For the Traffic lights, sign boards, Flags, zebra crossing, cars and trees the detection accuracy is greater than 70% because these objects mostly are with regular shape with specific designs, the user will face these kinds of obstacles with high probability. Other objects such as Roads and walls detection accuracy is medium because of their complicated structures. With the lowest detection accuracy, the objects are humans because of their complicated structure and high range of matching key points, for which a lot of data needs to be trained. The details of the object accuracy and errors detection is given in table 1 below.

Table 7. 1: Object detection accuracy using SURF

Names of Obstacles	**Detection**	**Error**	**Accuracy**
Cars	High	20%	80 %
Humans	Low	60%	40 %
Traffic Lights	High	0 %	100 %
Zebra Crossing	High	10%	90 %
Roads	Medium	40%	60 %
Walls	Medium	40%	60%
Buildings	Hight	30%	70%
Flags	High	10%	90%
Sign Boards	High	0%	100 %
Trees	High	20%	80 %
Maps	Medium	25%	75 %

The Signboard, parking, entrance gates, cars, buildings, flags, people, trees, zebra crossing, and roads are main objects detected from different scene to help the sight impaired people in this project, among them some are the main obstacles which most probably user faced everywhere. All the obstacles and objects were accurately detected shown in below figure 7.1, the project successfully provide the demo for sight impaired people from entrance gates to all other barriers which the user can face during walking speeds. The differences in object detection depends on the trained objects, the more data provides to the system the better the result will be for the user. The more trained objects gave the prototype more better results for processing and, increasing the rates of successful obstacles detection from the scene. Irregular objects and obstacles can decrease the successful rate of the demo. The accurate obstacle detection was possible with all known objects which normally occurred during the path and roads. The traffic lights, sign boards, entrance gates, poles, and trees had a very successful detection rate when compared to large obstacles such as the wall.

The car is also considered an obstacle of irregular shape, but the prototype correctly identifies it. This shows that the size or large surface will affect the speed of successful recognition. Traffic cones and masts have a lower detection rate than major obstacles such as walls. In one scene many obstacles were successfully detected. In addition, no false detection or detection was detected. The size and surface of traffic cones and masts are considered small, which may affect the accuracy of successful identification. The successful detection rate of irregular shape objects is different from other obstacles. Obstacles can be detected correctly form normal distance, long distance can be a problem for the analyzing of objects from the scene. A possible explanation for this is the height of the two barriers. When the user approaches a shorter object, the exposed surface is reduced, resulting in a lower rate of successful detection.
The comparison of SURF algorithm with other techniques are given below in table 2, the performance results for different image key points matches with accuracy is listed in table2. The results are very clear that among all other matches, SURF algorithm provides the best results for all the given images with respect.

Table 7.2: Comparison of Object detection Algorithm with SURF

Comparison of Performance Analysis of SURF Algorithm			
Algorithms	**Total Time (sec) Of Matching**	**Accuracy**	**Number of Correct Matches**
SURF	79135.063	%89.54	696.295
ORB	35074.987	%63.78	1414.503
SIFT	196487.667	%68.82	280.432
BRISK	35133.182	%63.52	1275.433

Overall the aim of the project was achieved. The objects of interest were successfully detected from the given scene images. After detection, the objects were bordered with red color in the final scene images. The final scene image also included, on the top, its GEO coordinates.

The final results are shown in figure 7.2.

Figure 7.1: Results

8. CONCLUSION

The main aim of the project work is to provide such a system which can assist the VI in traveling. The system should be able to detect the objects of interest, so that thy can be avoided and show GEO coordinates on the different scene images. The simulation was successfully achieved. Both mentioned tasks at hand were successfully addressed in this project work. The system took pre-taken scene images and performed object detection. These detected objects were highlighted on the original scene image. The results contain the scene image with highlighted objects and geo coordinates. This simulation can provide a new alternate option for the VI to utilize. The can also be customized according to each individual user.

The project shows that how the detection of objects can be converted from a relatively slow R-CNN algorithm to current optimization methods such as SURF, BLOB analysis and SVM algorithms. This development is largely independent of the structure of the classic network. Instead, it refers to the use of the convolution network and the calculations that occurred before and after the classic network. In the previous approach, there were several separate phases, including pre-treatment, zone formation, calculation of the completely bound layer, and final classification. In recent approaches such as SURF, which are used in this project, which focuses on key points or matching center points, these steps include convolutional networks while maintaining the integrity of the basic SVM model. After the results, the winner of the 2014 ImageNet competition for Object detection [26] has once again become a model for many independent components.

People with visual impairments have difficulty moving easily from one location to another. To move safely, you usually need support, for example: B. with a stick, a guide dog or a human friend. These options are limited, for example, a stick device cannot detect a hanging obstacle, in case of the helper dog has a limited life and people cannot always accompany people with visual disabilities. There are limitations in literature research, guidance and mobility professionals, and informal observations and discussions that are not based on views of existing barriers. In fact, researchers around the world have designed and developed numerous applications aids to help people with visual impairments. However, there were not enough tools to identify obstacles that could help people with visual impairments to avoid collisions. Many prototype barrier detectors have been developed, but they are bulky or obstructive to be less likely in everyday life. In addition, the performance of the prototype has not been sufficiently evaluated. The result of the study showed that it requires a low cost, a low calculation effort, a small compact size, a lightweight, portable and non-blocking obstacle detector to solve the problems faced by people with visual impairments.

8.1. FUTURE WORK

Currently the project is a simulation and is satisfying proof of concept. It is not in the form of launched as a product. Future work that can be done is, acquire the scene's in real time and perform the tasks achieved in this project work. MATLAB can be replaced by a hardware circuit that can detect the object and help the VI to avoid them using voice commands.

A smart phone can be used as a whole end solution. Scenes can be taken in real-time. After which the objects can be detected and avoided using voice comands. Goggle maps can be integrated in to the application so that the VI can travel any where without any predefined path.

REFERENCES

[1] Wikipedia: Visual Impairment. Accessed on: May. 20, 2017.[Online]. Available: https://en.wikipedia.org/wiki/Visual_impairment

[2] Difference between | Descriptive analysis and comparisons: Difference between Blind and Visually Impaired. Accessed on: May. 20, 2017. [Online]. Available: http://www.differencebetween.info/difference-between-blind-and-visually-impaired

[3] Medicine net: Blindness. Accessed on: May. 20, 2017. [Online]. Available: http://www.medicinenet.com/blindness/article.htm

[4] U.S national library of Medicine: Visual acuity. Accessed on: June. 8, 2017. [Online]. Available: https://medlineplus.gov/ency/article/003396.htm

[5] NCBI: GBD 2015 Disease and Injury Incidence and Prevalence Collaborators. Accessed on: June. 8, 2017. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/27733282

[6] WHO: Visual impairment and blindness. Accessed on: May. 8, 2019. [Online]. Available: http://www.who.int/mediacentre/factsheets/fs282/en/

[7] University of Tartu: Digital image processing. Accessed on: Jan. 5, 2019. [Online]. Available: https://sisu.ut.ee/imageprocessing/book/1

[8] Wikipedia: Computer. Accessed on: Jan. 5, 2019. [Online]. Available: https://en.wikipedia.org/wiki/Computer_vision

[9] Dr. Reinhard Klette: Concise computer vision, Springer-Verlag London, 2014

[10] H. Golnabi and A. Asadpour: Design and application of industrial machine vision systems, Robotics and Computer-Integrated Manufacturing, vol. 23, no. 6, pp. 630–637, 2007.

[11] T. Morris: Computer vision and image processing, Palgrave Macmillan, 2004.

[12] M. Schaeferling, U. Hornung, and G. Kiefer: Object recognition and pose estimation on embedded hardware, Surf-based system designs accelerated by FPGA logic, International Journal of Reconfigurable Computing, vol. 2012, Article ID 368351, 16 pages, 2012.

[13] Math Works: Object detection in computer vision. Accessed on: Jan. 5, 2019. [Online]. Available: https://www.mathworks.com/discovery/object-detection.html

[14] D. G. Lowe: Distinctive image features from scale-invariant key points, International journal of computer vision, vol. 60, no. 2, pp. 91–110, 2004.

[15] Antonio Ezio Frascarelli: Object detection for high productivity, 2015

[16] L. Iocchi, E. Menegatti, A. Bonarini, M. Matteucci, E. Pagello, L. C. Aiello, D. Nardi, F. Mastrogiovanni, A. Sgorbissa and R. Zaccaria et al: Development of intelligent service robots, Intelligenza Artificiale, vol. 7, no. 2, pp. 139–152, 2013.

[17] Tarun Agarwal: Obstacle avoidance robotic vehicle using ultrasonic for Obstacle Detection. Accessed on: Jan. 15, 2019. [Online]. Available: https://www.elprocus.com/obstacle-avoidance-robotic-vehicle/

[18] RLabayrade, D. Aubert and J. Tarel: Real Time Obstacle Detection in Stereovision on Non-Flat Road Geometry through v-Disparity Representation, Proc. IEEE Intelligent Vehicles Symp., vol. 2, pp. 17-21, 2002.

[19] Muhammad Asif: Smart shoe, Iqra University 2014

[20] Tam Nguyen, Josh Slonaker and Mohammed Kadous: Semi-Autonomous Wireless Controlled Robot Capstone, 2012

[21] Duy-Nguyen T, Wei-Chao Chen, Natasha Gelfand and Kari Pulli: SURFTrac: Efficient tracking and continuous object recognition using local feature descriptors, CVPR, pages 2937-2944, 2009.

[22] F. Ren, J. Huang, R. Jiang, R. Klette: General traffic sign recognition by feature matching, Proc. 24th Int. Conf. IVCNZ, pp. 409-414, 2009

[23] MathWorks: Image Processing toolbox. Accessed on: March.3, 2019. [Online]. Available: https://www.mathworks.com/products/image.html

[24] Quora: Which is better for more efficient and quicker Image processing-MATLAB, OpenCV or Octave. Accessed on: Feb. 7, 2019. [Online]. Available: https://www.quora.com/Which-is-better-for-more-efficient-and-quicker-Image-Processing-MATLAB-OpenCV-or-Octave

[25] Wikipedia: Speeded up robust features. Accessed on: Feb. 10, 2019. [Online]. Available: https://en.wikipedia.org/wiki/Speeded_up_robust_features

[26] Imagenet large scale visual recognition challenge 2016. Accessed on: March. 18, 2019. [Online]. Available: http:// image-net.org/challenges/LSVRC/2016.

[27] Miah , Md. Razu: A unique Smart Eye Glass For Visually Impaired People 2018, 10. 1109/ICAEEE.2018.8643011

[28] M. Crawford, *Catching the Sun*, American Society of Mechanical Engineers, Feb. 2013. Accessed on: Nov. 2, 2017. [Online]. Available: https://www.asme.org/engineering-topics/articles/renewable-energy/catching-the-sun