DRISTI: Dynamic Ranging by Image Segmentation and Terrain Imaging

Samanth S Mokshagundam, Suhas G, Suhas T Shanbhogue, Suraj S
Department of Electronics and Communication Engineering, BMSCE

Under the Guidance of
Dr. B. S. Nagabhushana
Professor
Department of Electronics and Communication Engineering, BMSCE

***Abstract*—Vision is the single most point of information input for a multitude of animals including humans. But it is estimated that 253 million people are visually challenged and 36 million people are blind. These people are faced with a multitude of challenges even in completing normal tasks in life. They are dependent on other factors or people for daily life. The advancement of science and technology in the recent years has enabled significant progress in providing support to disabled people. This project is one such attempt to improve the quality of life for the visually challenged by restoring the ability to self-navigate. It attempts to provide an effective and efficient platform to enhance the perspective of the surrounding to the visually challenged. This is achieved by design of an obstacle detecting system based on disparity maps to perceive the proximity of the obstacle. The implementation consists of a Raspberry Pi to capture images and give feedback to the user. Raspberry Pi is connected locally to a laptop wirelessly to send images and then processed on laptop to find obstacles. Feedback control signal is sent back to the Raspberry Pi. According to this, suitable audio feedback is given to the user.**

Keywords: *Visually Challenged, Obstacle detection, Stereo Vision, Disparity, Audio Warning.*

I. Introduction

A. Preface

AUtonomous navigation is an extremely important factor for people ailing from visual impairment. The traditional and common mobility aids for such people have been walking cane (also called white cane or stick) and guide dogs. The major drawbacks of these are obstacle range coverage and high maintenance respectively. With the rapid development in technology in computing, network, and sensors, pervasive technology has achieved significant progress and is being put into use gradually. This has opened up new avenues and potential for intelligent navigation capabilities. Past decade has brought a lot of interesting products as Electronic Travel Aids devised to help the blind people to navigate safely and independently as much as possible. Different possible approaches for Obstacle Detection include Vision, Ultrasound, Laser and RADAR, for Warning system, tactile and acoustic. In this project we implement a stereo vision based approach to detect obstacles, and an acoustic warning system. Situations like holes and stairs are not considered specifically in this project. The project is implemented on a raspberry pi connected locally to a laptop via Wi-Fi.

Though Stereo Analysis is a powerful tool, its computational complexity has necessitated the use of processing on a laptop, within a local network with the Raspberry Pi. Being a near real-time system and having used depth estimation and segmentation, our project is aptly named DRISTI: Dynamic Ranging by Image Segmentation and Terrain Imaging.

B. Problem Definition

Let's look at an ordinary day with the eyes of a blind person. There is so much to take for granted which is not given to others, for example, because they can't see. What happens with people who are totally blind? All too frequently, blindness affects a person's ability to self -navigate outside well known environments and even simply walking down a crowded street. Day-to-day things are done very differently and they are dependent for others for their basic tasks such as organizing toiletries, separating medicines etc. Sometimes, accessible technologies will solve the problem. This includes any technology which can be used equally well by those who can see and those who can't. In other cases, special technologies need to be used to achieve the same that others do just by using their sight. Dynamic Ranging by Image Segmentation and Terrain Imaging (DRISTI), is one such project specially designed which aims at assisting the visually challenged for navigation in simple environment. With DRISTI we aim to equip the blind to detect obstacles. They will be able to gauge the proximity between themselves and the object.

C. Objectives

The objectives of the project can be subdivided as:

- Obstacle detection: Segregate the foreground and the background. By utilizing the principle of depth estimation based on stereo imaging, the foreground and background of a given scene are segregated. A single scene is captured in two different perspectives with the help of two cameras. The shift in perspective of the two images is estimated. Based on this estimation, foreground and background can be separated because the foreground will have a relatively larger shift when compared to the background.

- Localization of obstacles: Determines the region where the object is present. After determining the presence of an obstacle, the region where the obstacle is present is determined. This is a very crucial step for giving feedback to the user.
- Warning system: After the region of an obstacle is determined, appropriate feedback must be given to the user. We plan on providing audio feedback to the user purely because it is the most primary sense a visually challenged individual can possess. Instead of utilizing a human voice for feedback, we plan on providing an intuitive acoustic feedback consisting of tones of varying frequencies.

II. LITEARTURE SURVEY

In literature, there are many different approaches for obstacle detection and warning system, such as:

- For obstacle detection: Vision, Ultrasound, Laser, and RADAR or a fusion of these.
- For warning system: Acoustic (TTS, Sound), Tactile (TDU, Vibratory motors etc.)

Below are some of the most relevant to our approach:

In [1], authors have used ground-plane detection algorithm using RANSAC and obstacles detected on this ground plane is intimated to the user via acoustic feedback. A ground plane estimation algorithm based on RANSAC plus filtering techniques allows the robust detection of the ground in every frame. A polar grid representation is proposed to account for the potential obstacles in the scene. The design is completed with acoustic feedback to assist visually challenged users while approaching obstacles. Beep sounds with different frequencies and repetitions inform the user about the presence of obstacles. Audio bone conducting technology is employed to play these sounds without interrupting the visually challenged user from hearing other important sounds from its local environment.

In [2], Obstacle detection is achieved by using Optical Flow with Stereo Vision which gives good results even for moving obstacles, unlike the previous one which is more concerned with static results. Accurate detection of moving obstacles from a moving vehicle is at the core of safe autonomous driving research. Stereo vision based sensors have been extensively used for this task as they are passive and provide a large amount 3D and 2D data. However, since no motion information is revealed, in intersections or crowded urban areas, static and dynamic objects immediately next to each other, or closely positioned obstacles moving in different directions are often merged into a single obstacle leading to dangerous misinterpretations.

In [3], authors implement an autonomous automobile navigation based on obstacle detection using Unscented Kalman Filter. Its a very comprehensive detection system, but also requiring extensive computation which is suitable for an automobile but not for a wearable such as ours.

Reference [4], is another project which is aimed at autonomous navigation for automobiles. It computes a V-disparity map, from which obstacles are determined. Many roads are not totally planar and often present hills and valleys because of environment topography. Nevertheless the majority of existing techniques for - road obstacle detection using stereovision assumes that the road is planar. This can cause several issues: imprecision as regards the reel position of obstacles as well as false obstacle detection or obstacle detection failures. In order to increase the reliability of the obstacle detection process, this paper proposes an original, fast and robust method for detecting the obstacles without using the flat-earth geometry assumption; this method is able to cope with uphill and downhill gradients as well as dynamic pitching of the vehicle. Our approach is based on construction and investigation of the ”v-disparity”' image which provides a good representation of the geometric content of the road scene.

The authors of [5] use a similar approach as us, using a Kinect device to obtain a RGBD image which is used to detect pre-defined objects and this is related to the user using Tongue Electrode matrix. This system consists of two main components: environment information acquisition and analysis and information representation. The first component aims at capturing the environment by using a mobile Kinect and analysing it in order to detect the predefined obstacles for visually challenged people, while the second component tries to represent obstacles information in the form of electrode matrix.

The authors of [6] describe obstacle detection and mapping technique for mobile robot, based on depth matrix and drop-off detection using a laser based 3-DOF local SLAM. A mobile robot operating in an urban environment has to navigate around obstacles and hazards. Though a significant amount of work has been done on detecting obstacles, not much attention has been given to the detection of drop-offs, e.g., sidewalk curbs, downward stairs, and other hazards where an error could lead to disastrous consequences. This paper proposes algorithms for detecting both obstacles and drop-offs (also called negative obstacles) in an urban setting using stereo vision and motion cues. They propose a global colour segmentation stereo method and compare its performance at detecting hazards against prior work using a local correlation stereo method they implemented a U-disparity map to find the depth and applied a modified particle filter to detect multiple obstacles and track them. Vision systems provide a large functional spectrum for perception applications and, in recent years, they have demonstrated to be essential in the development of Advanced Driver Assistance Systems (ADAS) and Autonomous Vehicles. In this context, this paper presents an on-road objects detection approach using a new strategy in obstacle extraction from U-disparity

One more [8], approach for Unmanned Ground Vehicles using a Horopter-based Stereo vision and terrain imaging by 3D model formed based on real-time video processing. Sarnoff's next-generation video processor, the PVT-200, is used to demonstrate real-time algorithms for stereo processing, obstacle detection, and terrain estimation from stereo cameras mounted on a moving vehicle. Sarnoff's stereo processing and obstacle detection capabilities are currently being used in several Unmanned Ground Vehicle (UGV) programs, including MDARS-E and DEMO III. Sarnoff's terrain estimation capabilities are founded on a "model-based directed stereo" approach. A collaborative research is demonstrated between Sarnoff and Universitat der Bundeswehr Miinchen, where we are studying vision processing for autonomous off-road navigation as part of the AUTONAVprogram.

After studying the above papers and their relevance to our objective, we decided to take a Stereo Vision Disparity approach to implement obstacle segregation based on their proximity to the user. We also used an audio feedback technique based on beeps instead of words, as it would feel more intuitive to the user.

III. Hardware Setup

The hardware setup of our project comprised of:

- Raspberry Pi
- Stereo Camera Setup c170 webcams $\times$ 2
- Earphone
- Laptop

Fig. 1. Hardware Setup

IV. Implementation and Methodology

A. Block Diagram

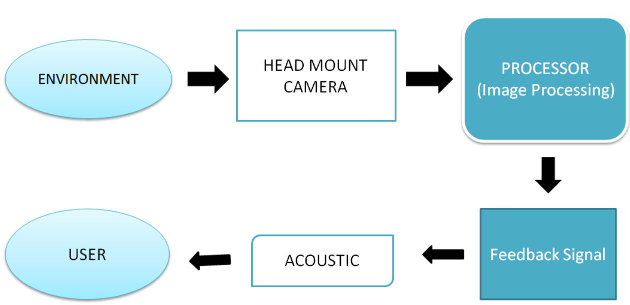

Fig. 2. Block Diagram

With the help of two Logitech C170 web-cameras, two different perspectives of the same scene are captured. Several image processing techniques such as Rectification, Disparity Computation, Segmentation are then applied to the image pair obtained by the processor to identify the relative proximities of potential obstacles in the captured scene. The proximities of the obstacles are then appropriately classified and based on the region of occurrence of the obstacles, audio feedback are given to the user which help in autonomous navigation.

B. Project Flow

Our project can be subdivided into 2 flows:

- Pre-processing.
- Implementation Process.

Pre-processing:

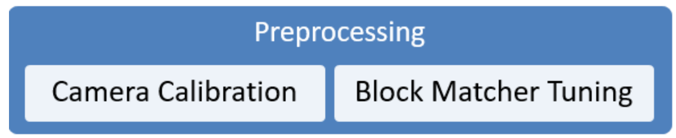

Fig. 3. Pre-processing

Pre-processing involves 2 steps:

- Camera Calibration:
  The Logitech C170 web cameras predominantly suffer from two types of distortions: Radial distortion in which straight lines will appear curved and tangential distortion which occurs because image taking lenses are not aligned perfectly parallel to the imaging plane. So some areas in image may look nearer than expected. In order to remove these undesired irregularities from the two images, the individual camera parameters are required. The camera properties include the camera matrix, distortion coefficients, Rotation and Translation vectors, Projection matrix etc. To find all these parameters, we provided 100 images of a 7X10 chessboard in different orientations. One such orientation is shown below. We find some specific points in it (square corners in chess board). We know its coordinates in real world space and we know its coordinates in image. With these data, a mapping of object points to image points is done. This procedure is applied to a Stereo Image pair. When corresponding points in a set of images is known, it's possible to find the relative translation vector and rotation vector, which gives rise to the distortion. Hence now we know all the parameters required to rectify the image.
- Block Matcher Tuning
  Semi Global Matching is a method which we are using to compute disparity. But it has a lot of parameters to be tuned for a particular hardware setup. So, to help in tuning we wrote an application to tune the parameters on the fly while observing the disparity.
  Parameters to be tuned were:
  - minDisparity - Minimum possible disparity value. Normally, it is zero but sometimes rectification algorithms can shift images, so this parameter needs to be adjusted accordingly.

Fig. 4. Image from the left and right cameras

– numDisparities - Maximum disparity minus minimum disparity. The value is always greater than zero. In the current implementation, this parameter must be divisible by 16
– blockSize - Matched block size. It must be an odd number $\geq$1. Normally, it should be somewhere in the 3...11 range.
– P1 - The first parameter controlling the disparity smoothness
– P2 - The second parameter controlling the disparity smoothness. The larger the values are, the smoother the disparity is. P1 is the penalty on the disparity change by plus or minus 1 between neighbor pixels. P2 is the penalty on the disparity change by more than 1 between neighbor pixels. The algorithm requires P2 ¿ P1
– disp12MaxDiff - Maximum allowed difference (in integer pixel units) in the left-right disparity check. Set it to a non-positive value to disable the check.
– preFilterCap - Truncation value for the prefiltered image pixels. The algorithm first computes x-derivative at each pixel and clips its value by [-preFilterCap, preFilterCap] interval. The result values are passed to the Birchfield-Tomasi pixel cost function.
– uniquenessRatio - Margin in percentage by which the best (minimum) computed cost function value should ”win” the second best value to consider the found match correct. Normally, a value within the 5-15 range is good enough.
– speckleWindowSize - Maximum size of smooth disparity regions to consider their noise speckles and invalidate. Set it to 0 to disable speckle filtering. Otherwise, set it somewhere in the 50-200 range.
– speckleRange - Maximum disparity variation within each connected component. If you do speckle filtering, set the parameter to a positive value, it will be implicitly multiplied by 16. Normally, 1 or 2 is good enough.

A screenshot of the application is given below:

The importance of parameter tuning can be seen by the difference in disparity images obtained:

C. Implementation

- Image Capture

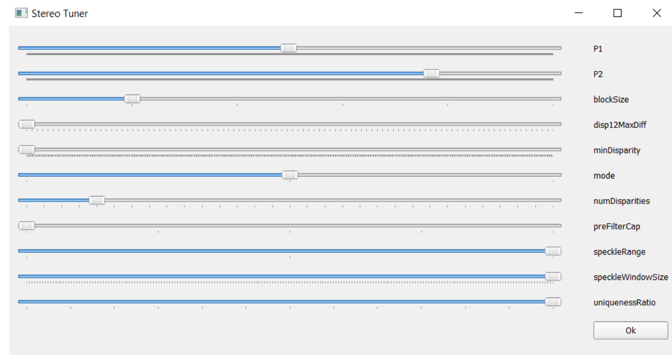

Fig. 5. Block matcher Tuning Application

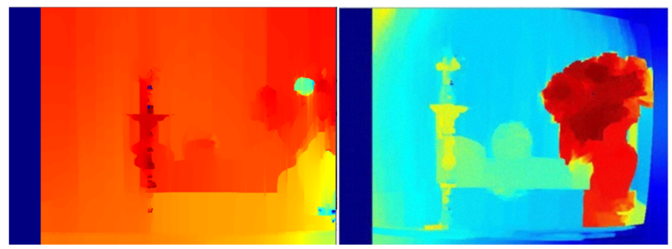

Fig. 6. Comparision before and after Tuning

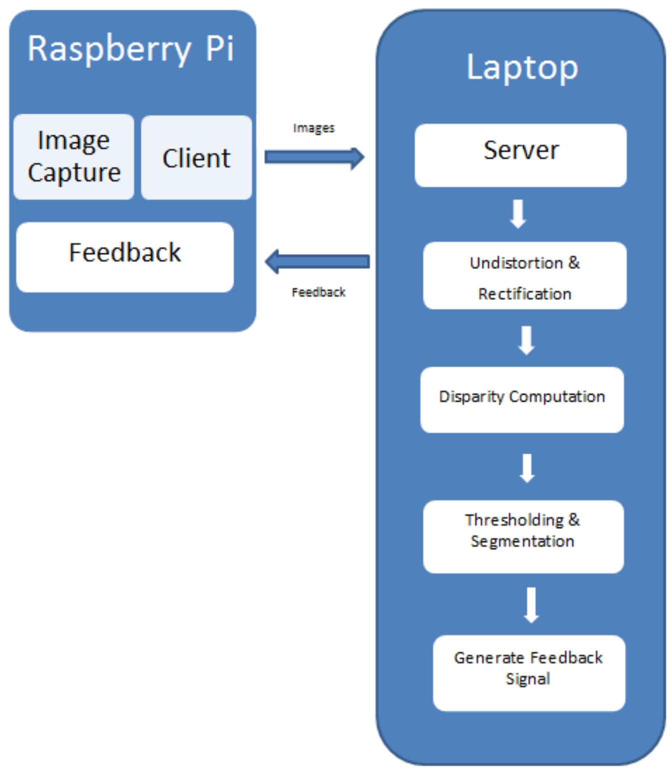

Fig. 7. Implementation Flow

Using two Logitech C170 web cameras, two images each representing different perspectives of the same scene are captured. For the purpose of illustration of the concepts detailed below, we make use of a stereo image pair we have captured with the help of the web cameras we have used. The stereo pair is shown below.

Fig. 8. Stereo Pair

- Client.
  We are using Raspberry Pi as the client to capture image and then convert the image to grey-scale and send it to the laptop over wireless LAN. Raspberry Pi 3 Model B is having on board BCM43438 wireless LAN, which we can use to connect to a mobile hotspot, also to which the laptop is connected as well. Since this is a real-time application, streaming protocol based on UDP is used.

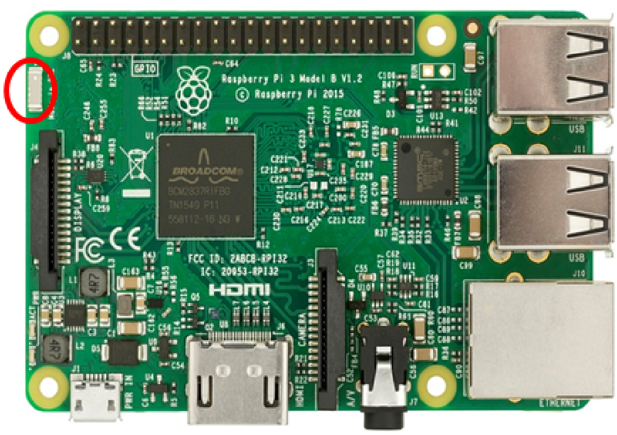

Fig. 9. Antenna Module on Raspberry Pi

- Laptop.
  Laptop acts as the server which receives the image from the raspberry pi and processes on it. After processing, laptop sends the feedback control signal.
- Undistortion And Rectification
  With the help of the camera parameters, the images captured from both the left and right cameras are subjected to image undistortion. The undistorted image is then subjected to rectification, where the image planes of both the captured images are made parallel to each other. Rectified image pair for the above test images is shown below.

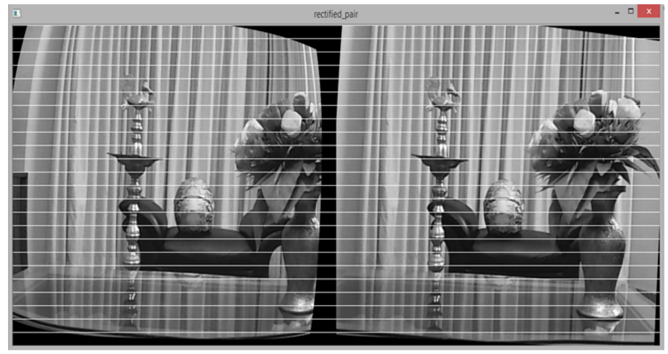

Fig. 10. Rectified stereo Pair

- Disparity Computation
  The two rectified images are now used to compute the disparity corresponding to both the images. Disparity is formally defined as the differences in x-coordinates on the image planes of the same feature viewed in the left and right cameras: xleft xright. The two algorithms which are widely used for disparity computation are BM (Block Matching) and SGBM (Semi Global Block Matching). We have employed SGBM for disparity computation because SGBM offers relatively more accuracy and clarity. The disparity in both grey scale and RGB is shown below.

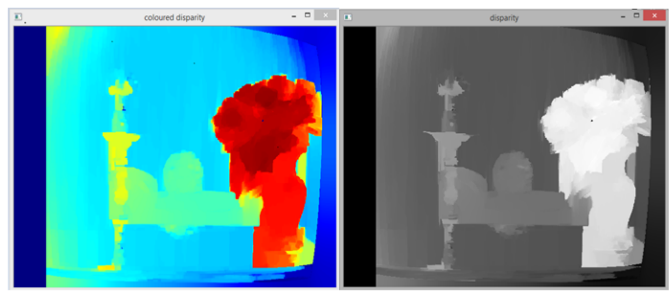

Fig. 11. Disparity - Grayscale and False Coloured

- Disparity Thresholding And Segmentation
  The disparity obtained is then subjected to 2-level binary thresholding indicating two levels of proximities. The proximity range for dividing the objects into these two regions is determined by the disparity intensity value. Each of the proximity image planes obtained are then subjected to segmentation. This is done by dividing each of these image planes into appropriate number of subdivisions. Presence of objects in these subdivisions is determined by contouring. The two regions of proximities are shown below

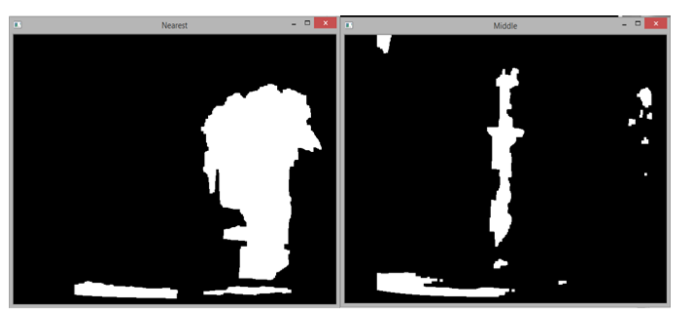

Fig. 12. Thresholding and Segmentation of Disparity - Closest and Farther

- Generating Feedback Control Signals.
  Based on the results of segmentation, appropriate audio signals are generated to indicate obstacles present in the

”Near” region .For this, 3 frequencies are taken and the concept of stereo audio is used to differentiate the position of obstacles.The truth table for audio feedback delivery is given in the table below.

TABLE I
FEEDBACK GENERATION TRUTH TABLE

IMAGE REGIONS			OUTPUT
LEFT	CENTER	RIGHT	
0	0	0	No Feedback
0	0	1	F1, Right Earphone
0	1	0	F1, Both Earphone
1	0	0	F1, Right Earphone
1	1	0	F2, Left Earphone
0	1	1	F2, Right Earphone
1	0	1	F2, Both Earphone
1	1	1	F3, Both Earphone

- Feedback
  Based upon the feedback control signal, appropriate frequencies are generated and given to the user through earphones. Here the concept of Stereo sound is used.

V. RESULTS AND DISCUSSION

A. Results

Project DRISTI works as proposed; two cameras are used to obtain live visual feed of the surroundings of a blind person. This is given as input to the Raspberry Pi, which computes the grey scale image and sends it to the laptop over a Local Area Network for further computation. The laptop is used to calculate disparity from the two camera using SGBM algorithm. When the blind person encounters an object the disparity calculated by the laptop will be utilised by the Raspberry Pi to give an acoustic feedback based on the proximity of the said object. Depending upon which ear gets the audio signal and also upon the frequency of the signal the number of objects and the region/s in which the object is present can be determined. The user can use this data to avoid the obstacle.

B. Discussion

In this project, visual assistance system for the blind, based on stereo imaging was proposed. Since the traditional methods such as walking stick and guide dog are not efficient and reliable, there was a need for new technology which overcomes these drawbacks. For the purpose of assisting the visually challenged, in order to gauge the distance, stereo imaging technique has been utilised. It is advantageous over other techniques which use sensors for measuring distance. E.g.: in the case of ultrasonic sensors objects need to be present in the line of sight of the sensor, In addition they are not able to distinguish between a single object and multiple objects and also the area in which the object is present in. In order to determine these using a sensor like ultrasonic sensor, multiple sensors along many directions will be required which will increase the cost significantly and the arrangement of sensors will again not be feasible, so the best approach would be to use two cameras and obtain stereo images for object detection as well as depth mapping of the objects. For every scene two images were obtained using the cameras. In order to obtain the depth mapping of the scene, different algorithms were tested upon the images to determine the best possible algorithm as shown in Fig. 13 and Fig. 14.

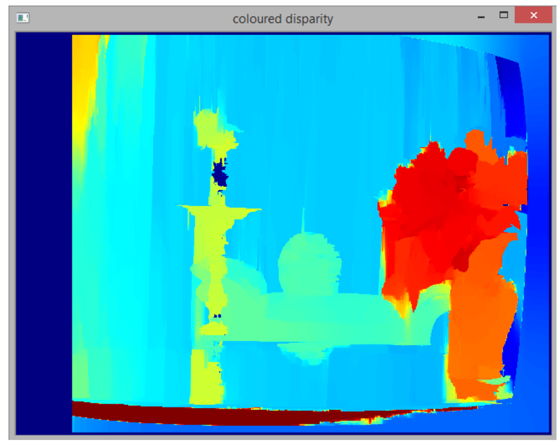

Fig. 13. Block matcher disparity

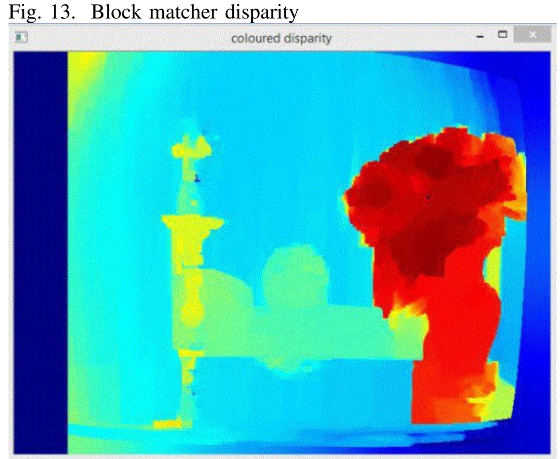

Fig. 14. Semi-Global Block matcher disparity

From the outputs for the sample set presented above, we can conclude that SGBM provides a smoother image with fewer distortions. For this reason SGBM algorithm was chosen in order to calculate the disparity. Technology is to be used for the welfare of all mankind, we would like to equip all the visually challenged individuals with this new tool for their navigational aid, but the cost of the project makes it less affordable for some, However with proper government funding this can be overcome. Project DRISTI will impede with the normal ear function as the user will have to wear earphones, but the new research field which uses bone conduction technology will be able to liberate the ear. Apart from these shortcomings of the system there are no major flaws that may give us a thought to drop the idea or lower our interests in such systems. The perks and gains from this system are far beyond their flaws.

The benefits of the proposed system are:

- Obstacle recognition at a distance: One of the major advantages is that the blind person will be able to identify the presence of an object from a distance without having to come in contact with the obstacle.
- Reduced human error: as the primitive walking stick is replaced with more accurate technology there is reduction in error, both false positives and false negatives, for obstacle detection.
- Ease of navigation: Indoor navigation as well as outdoor navigation is made easier.
- Improves the self-esteem of the blind as they become more independent in their navigation and day to day activities.
- Saves time: Navigation time from one place to another is reduced with the help of this tool.

Since the orientation of the object with respect to themselves is known, they will be able to determine the necessary course to be taken in order to avoid the obstacle.

VI. Conclusions and Future Work

A. Conclusions

At the end of this project, the conclusion drawn, was that a new technology of obstacle detection could be brought into practise using stereo imaging for helping the blind navigate through to their destination without having to take too much of a risk in reaching their destination. DRISTI (Dynamic Ranging by Image Segmentation & terrain Imaging), would be of good help to the blind for object detection and autonomous navigation. A noteworthy advantage of this project is that it does not depend on which type of obstacle (vehicle/human/animal, etc.) is approaching towards the visually challenged person or is being approached to by the person, due to the feedback, the person knows about the danger ahead and is able to avoid collision with these obstacle thereby saving himself/herself from any sort of injury/damage that could be caused otherwise. DRISTI also provides hands free navigation to the blind and is a reliable means of feedback system to ensure that they are well informed of any obstacles ahead. The person does not have to constantly try to gauge for obstacles with the help of any aid such as walking stick, cane, etc. There also will not be any restrictions in any well-lit environments for the blind to have this device help them unlike other aids such as guide dogs and won't get stuck in pavements, pits, etc. like walking sticks & canes. The feedback is reliable as it not only helps detect obstacle, but also helps to know in which direction the obstacle ,depending upon from which direction(left or right the audio feedback is obtained and helps to also get a gauge of the approximate distance of the obstacle depending on the intensity of the feedback. Hence it is clear that autonomous navigation through audio feedback would be of great help to the blind, especially in a country like India with the huge blind population.

B. Future Work

Our project mainly aimed at detection of obstacles and would mainly help in indoor navigation for the blind people and in roads without too much gradient change and potholes. Following are the advancements that can be made to enhance the perspective of the surroundings even better to the blind people so that they can further navigate without fear and any associated risk.

- Pothole detection: Pothole detection would be of great importance where maintainence of roads isn't done properly.
- Gradient detection or Stairacase Detection: Any sudden change in slope would be dangerous. Similarly, stairs though wouldn't count as obstacle might indeed be dangerous without proper warning.
- Obstacle Recognition: Recognition of object in front will lead to a greater perception of the environment.
- GPS Tracking System: A GPS tracking system would be extremely helpful in outdoor environments, imparting the ability to viusually impaired to navigate independently.

References

[1] Alberto Rodrguez, J. Javier Yebes, Pablo F. Alcantarilla, Luis M. Bergasa, Javier Almazan and Andrs Cela, "Assisting the Visually Impaired: Obstacle Detection and Warning System by Acoustic Feedback", Open Access, Sensors, ISSN 1424-8220

[2] Cosmin D. Pantilie, SilviuBota, Istvan Haller and SergiuNedevschi,"Real-time Obstacle Detection Using Dense Stereo Vision and Dense Optical Flow", 978-1-4244-8230-6/10/$26.00 2010 IEEE

[3] Qian Yu, HelderAraujo,Hong Wang, "A Stereovision Method for Obstacle Detection and Tracking in Non-Flat Urban Environments", Autonomous Robots 19, 141157, 2005 2005 Springer Science + Business Media, Inc. Manufactured in The Netherlands.

[4] Raphael Labayrade, Didier Aubert, Jean-Philippe Tarel, "Real Time Obstacle Detection in Stereovision on Non Flat Road Geometry Through "V-disparity" Representation".

[5] Van-Nam Hoang, Thanh-HuongNguyen,Thi-Lan Le, Thanh-Hai Tran, Tan-PhuVuong, Nicolas Vuillerme,"Obstacle detection and warning system for visually impaired people based on electrode matrix and mobile Kinect", Open Access at Springer 26 July 2016

[6] AniketMurarka, Mohan Sridharan, Benjamin Kuipers,"Detecting Obstacles and Drop-offs using Stereo and Motion Cues forSafe Local Motion",published at https://web.eecs.umich.edu/ kuipers/

[7] Bihao Wang, Sergio Alberto Rodriguez Florez, Vincent Fremont,"Multiple Obstacle Detection and Tracking using Stereo Vision: Application and Analysis", published in HAL archives, https://hal.archives-ouvertes.fr/hal-01095618

[8] Robert Mandelbaum, Luke McDowell, Luca Bogoni, Barry Reich, Mogens L. Hansen,"Real-Time Stereo Processing, Obstacle Detection,and Terrain Estimation from Vehicle MountedStereo Cameras",Proceedings of the 4th IEEE Workshop on the Applications of Computer Vision, October 1998, pages 288-289.