ODAV Literature Review

The world contains a total estimate of 2 million people that are visually impaired and unable to perform routine tasks that we might take for granted. One of the significant problems that is faced by the visually impaired is independent mobility. They are dependent largely on external help in the form of guide dogs, human support or white canes. However with the advancement of technology there have been recent attempts at solving this problem. A number of wearable object detection systems have been developed for assisting the blind people, however none of them have proven to be effective for mass scale manufacturing. These solutions were mainly focused around RFID, Kinect, sensor arrays and Stereoscopic technology.

1. Existing Technologies

1.1. RFID Technology

RFID stands for Radio Frequency Identification. It uses electromagnetic fields to automatically identify and track tags attached to objects. Each tag is given a unique ID. Its application in helping the blind is done in such a way that all obstacles in an environment are assigned their own tags. When a person is present in that environment, his/her distance from any object can be calculated using the RSSI (Received Signal Strength Indicator) values. However this solution is only feasible for a static environment where there aren't any moving objects. In addition to that any obstacles that do not have a tag will not be detected. Another major problem with solutions based on this technology is that it can only be deployed in a small environment, and large scale implementations are not feasible.

1.2. Kinect Technology

Kinect technology makes use of the patented Kinect technology that Microsoft developed for XBOX360. It makes use of camera based design in which a laser grid is projected through an infrared projector. A camera along with an inbuilt chip then ascertains the location a nearby object in 3D space. This method is called light coding. Using the Kinect we can obtain a disparity map of the surroundings and from that we can determine whether there are obstacles in the environment or not. The downside of this technology is that the Kinect must be in a static and stable position to ascertain the disparity map of the environment. In addition to that its range is questionable. Another disadvantage is that the hardware is fragile and must be handled with care, which makes it unsuitable for any wearable device. Power and processing requirements are also a concern for solution employing this technology.

1.3. Sensor Arrays

Sensor arrays and sensor based solutions offer a cheap alternative to other solutions however there results are not as good as other available technologies. Most of these solutions employ Ultrasonic sensors or Infrared sensors however both of these sensors cannot detect all kind of obstacles. For instance many attempts have been made at integrating the ultrasonic sensors and the walking sticks that the visually impaired already use to make a smart stick that detects nearby obstacles and gives feedback to the user. However the smart sticks fail to detect any obstacles that are hanging overhead and also there are issues with using ultrasonic sensors. Ultrasonic sensors work on the principle of detecting echo reflection and fail to detect a reflection if the obstacle's shape is curved as it will reflect the echo in a slightly different direction. In addition to that they do not work well in the presence of objects that can absorb ultrasonic signals such as sponges etc. Infrared sensors have their own issues and cannot be used in the presence of sunlight. In addition to that they do not work well in the presence of metallic objects that can absorb infrared light.

1.4. Stereoscopic Technology

Stereoscopic technology is based around using stereo cameras. They work on the same principle as the human eye, images are taken from 2 cameras placed at a small distance apart and then objects are identified in the images. After that their depth is perceived by generating disparity maps that allow us to convert 2D information from the pictures to 3D information regarding the depth of the objects in the images. Objects are detected via edge detection in this method. The disadvantages of this method is that it requires processing time and it fails to detect objects that do not have any distinguishable features when compared to the background. For instance it will not be able to detect any smooth objects or walls as there are no edges present. In addition to that good quality cameras are expensive and sensitive, therefore must be handled with care.

2. Existing Devices and Systems for navigation and object detection

2.1. Robot Sensing and Smart Phones for navigation

Robots need help to navigate through their surroundings which can be static and dynamic in nature. They employ stereoscopic techniques to navigate through unknown environments and the same principle can be used to help blind people navigate indoor and outdoor spaces independently as mentioned above.

One such remote sensing system is discussed in [1]. It was developed by E. Pissaloux along with his colleagues at the Institute of Intelligent Systems and Robotics at Pierre and Marie Curie University in France. The system consists of a pair of glasses equipped with cameras and sensors. The system produces a 3D map of the user's environment and their position is regularly updated and displayed in a tactile pad or an electronic portable Braille device. It was designed to allow blind people to navigate freely without any external aid such as walking cane etc.

2 cameras attached on either side of a pair of glasses are used to give a live feed to the processor, Both the cameras help the processor in generating a 3D image of the surrounding surface by generating a disparity map as explained in the portion of Stereoscopic technology. The system keeps track of the person's location and his/her speed by employing accelerometers and gyroscopes like the ones used in robot sensing. The person's location along with the disparity map are combined to determine the user's position in relation to other objects.

Another navigation is also explained in [1]. The system is software based and can predict the distance a robot has travelled using information obtained from its onboard sensors. This information can be used to track a person's movements based on their step length. The advantage of this system is that it is low cost. It was developed by E. Folmer and K. Bekris at the University of Nevada in Reno. It uses freely available 2D digital indoor maps, smart phones and inbuilt accelerometer and compass. Feedback to the user is given through artificial speech. The system would require to be trained according to each individual user first in order to give satisfactory results. However this system would only work in local environments and extending it to all environments would be unfeasible.

2.2. Real-time Localization in Outdoor Environments using Stereo Vision and GPS

For effective outdoor implementation of a navigation system for blind people that can also direct them to where they want to go, a lot of research has been done in attempts to integrate GPS into navigation systems. The GPS system takes in the location of a person and compares it with the digital maps stored in memory to find out the exact location of the user. The project in [2] is about localization of objects using stereo camera images with GPS sensor. The design enables obstacle detection over few meters along with the localization capabilities. The functioning of the system depends on the accuracy of the GPS systems and it is hard to use for areas where GPS is not available. Therefore it is not capable of operating in indoor environments.

2.3. Wearable Object Detection System for the Blind

It is not an easy task for visually impaired people to distinguish between different items, for examples packaged foods and medicine containers to them would appear the same as they would be relying only on their sense of touch. To resolve this problem RFID technology is used and every item is assigned its own tag. This technology can help the blind in their daily life activities. RFID technology has been implemented for many other applications as well such as scanning passports, highway toll collection etc. To improve mobility of the visually impaired as mentioned before, the RFID tags can be used in tandem with RSSI information to locate any nearby obstacles and inform the user of his/her distance from them.

2.4. Beacon Based Smart City for Visually Impaired

This was a project by Microsoft in which they employed Bluetooth beacons instead of RFID tags on obstacles. The working principle behind this was again RSSI and objects nearby to the user of the system would identify them through the unique tag assigned to them via their Bluetooth beacon. It was discontinued due to its non-feasible nature for implementation on a large scale. A small test area could be set-up however setting up the beacons for an entire city or even a block would result in large expenses.

2.5. Smart Canes for Obstacle Detection

In an effort to find a cheap solution to the problem of obstacle detection many attempts have been made to modify the walking cane that the blind already use to allow it to detect obstacles in front of it. Sensors are added to the cane so that it can detect objects in near vicinity however these solutions are not effective as the sensors do not always detect obstacles. Also these solutions require that the cane be close to the obstacle for it to be detected and thus mostly fail to detect hanging obstacles. This is why this is not a practically feasible solution.

3. Stereoscopic Technique

Stereoscopic technology as mentioned above is one of the commonly used methods for the purpose of navigation for robots and in some instances for blind people. Our project makes use of stereoscopic technique as well, therefore it is worthwhile mentioning a few technical details about this technique. There are mainly 2 methods used for measuring the distance of the object from the user; active and passive. In active techniques the distance is calculated by sending some signal to the object like radio signals, laser beams, ultrasonic waves etc. and measuring the time it takes for those signals to return. Passive techniques on the other hand use other methods for determining the distance. Stereoscopy is the most popular form of passive distance measuring technique by far. Our project makes use of both passive and active techniques. We are using ultrasonic sensors for the active method and stereo camera for the passive method.

The basic technique in passive method revolves around images taken from a stereo camera and then calculating the distance between the object and viewer. The formula for calculating the distance is given below:

$$D = \frac{Bx_0}{2\tan\left(\frac{\varphi_0}{2}\right)(x_L - x_D)}$$

Where:

- D = Distance between the object and camera
- B = Distance between the cameras
- X_o = Number of horizontal pixels
- Θ_o = The viewing angle of camera
- $(x_L – x_D)$ = The horizontal difference between the same object on both pictures

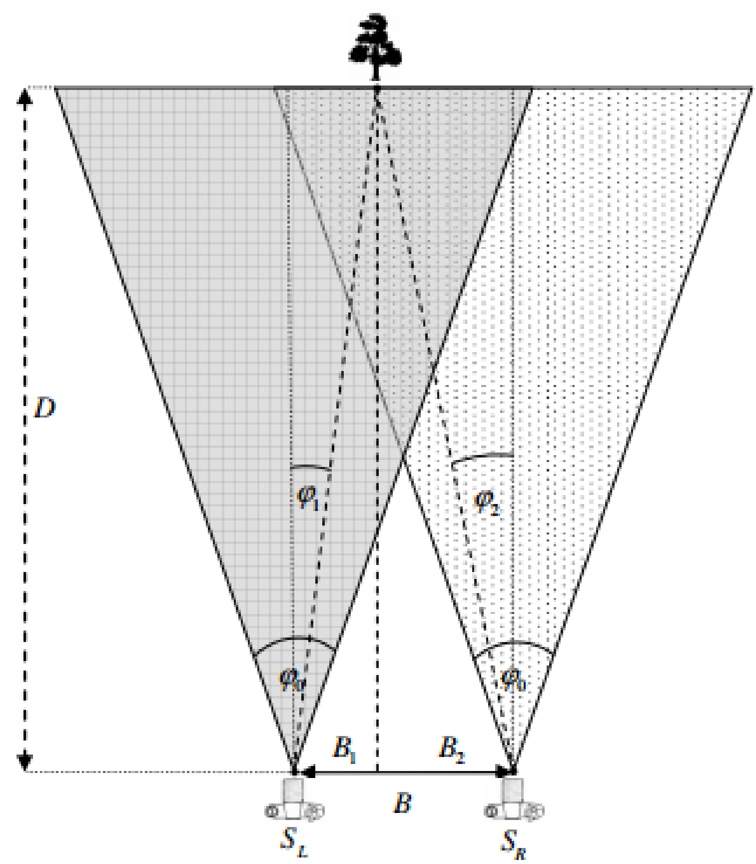

Figure 1: Picture taken with both cameras of a tree

Based on the above formula, disparity maps are generated, from which we can determine the distance of objects.

3.1. Issues faced in generation of disparity maps

The following issues arise and need to be dealt with before accurate disparity maps can be generated.

3.1.1. Barrel Distortion

Barrel Distortion occurs particularly in wide angle lenses, it is the phenomenon in which straight lines are curved in in the shape of a barrel. This phenomenon occurs because the size of image is smaller than the field of view of the camera, and therefore the image is squeezed in. If disparity maps are generated without resolving barrel distortion first then inaccurate depths are generated. Whilst generating disparity maps, it is our assumption that the camera lens is acting like a pinhole camera, which is by default planar in nature. Therefore we are assuming an epi-polar plane whereas in reality the camera is not planar in nature. Therefore it is imperative that we rectify the image to remove barrel distortion first.

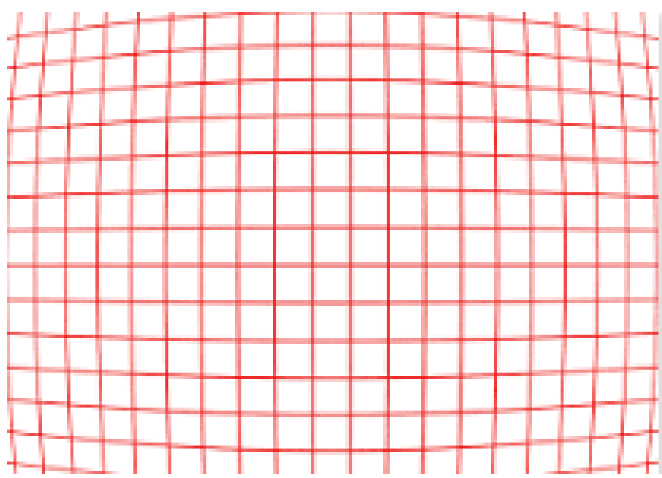

Figure 2 Barrel Distortion Effect

3.1.2. Radial Distortion

Radial Distortion is similar to Barrel Distortion and occurs in wide angle lenses. The only difference is that, in barrel distortion the image is squeezed in and as a result a barrel like protruding effect occurs in the image, however in radial distortion the exact opposite happens and a cavity effect occurs in the image. The effect of this distortion on the disparity maps is again false results.

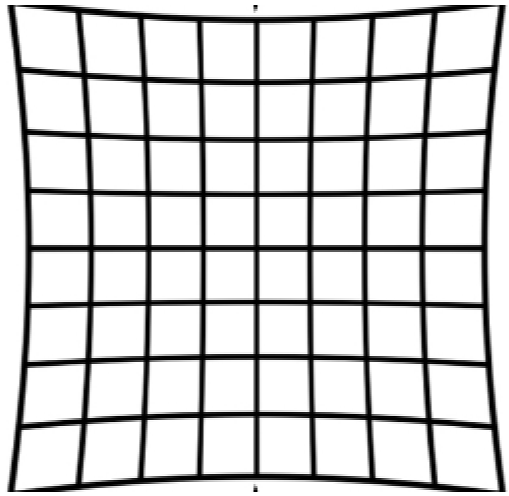

Figure 3 Radial Distortion Effect

3.2. Steps to taken to Resolve Issues

To resolve the issues of Barrel and Radial distortion, image rectification was done before calculating disparity maps. In order to do that however we required camera parameters which contain information such as focal length, Translation vector, Rotation Matrix etc. Therefore in order to obtain all this information we had to perform camera calibration, which yielded the above data. After calibration, we were able to rectify the images and obtain the correct disparity maps, on which the image processing was performed.

4. References:

[1] http://www.newscientist.com/article/mg21428625.700-robot-sensing-and-smartphones-helpblind-navigate.html

[2] Real-time Localization in Outdoor Environments using Stereo Vision and Inexpensive GPS