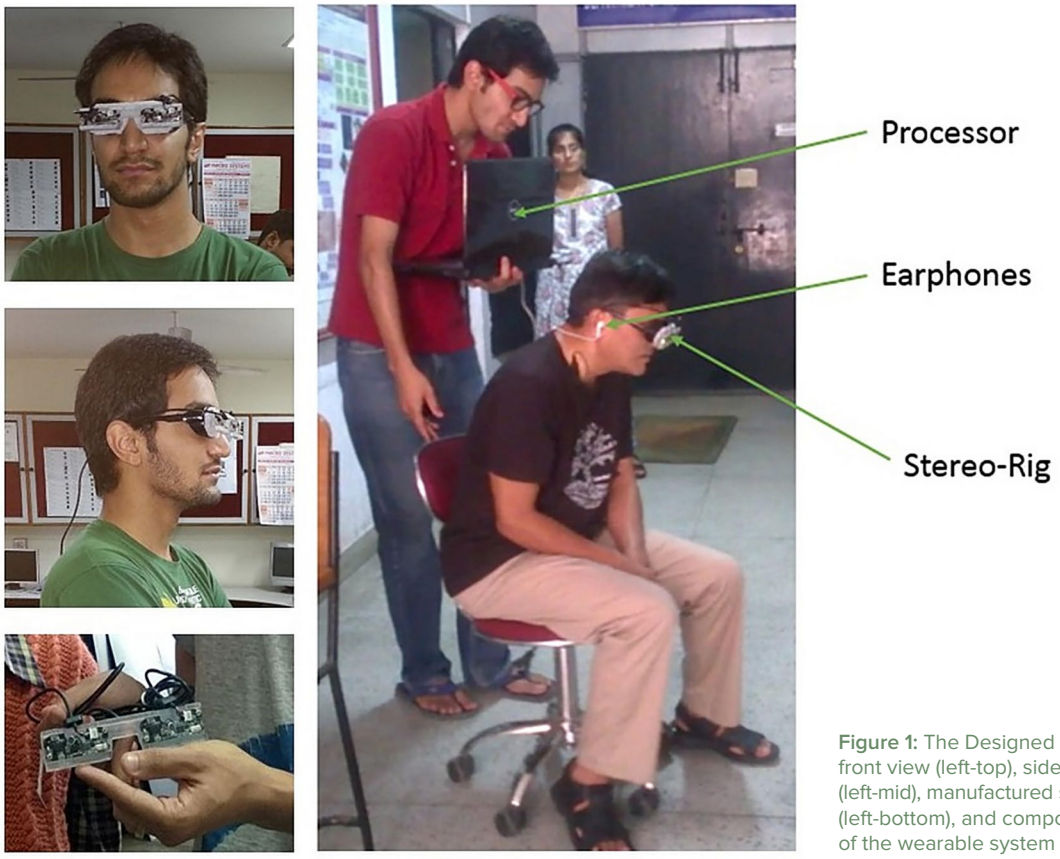

Figure 1: The Designed System—front view (left-top), side view (left-mid), manufactured stereo-rig (left-bottom), and components of the wearable system (right).

A Navigation System for the Visually Impaired Using Stereo Vision

Overview of the Project

In this work, we have successfully created a spectacle (**Figure 1** left-top and left-mid), which as proven by our experiments has the potential of replacing the cane used by visually impaired people.

The system exploits the use of two cameras fixed rigidly to one another (Stereo system) and a processing unit, which uses the algorithms from the field of stereo vision and robotics. The processor then gives the output through the earphones, in the form of an audio beep. In essence, the amplitude of the beep is inversely proportional to the distance of the closest obstacle in front of the user. This simplistic representation of space makes sure of 'zero information overloading' and at the same time proves as a useful device to the user. Since this is exactly based on the principle of the cane (stick used by the blind), the training requirements to use the spectacle are minimal.

The system is designed in a manner which not only solves the problem of obstacle avoidance, but also takes care of the ergonomic requirements, making the whole experience a joyful one. The simplicity, effectiveness and frugality of the system have been tested and verified by our experiments with blind-folded people.

Motivation and Need

Blindness prevails all around the world. In most cases people with visual impairment suffer from severe mobility issues, especially in countries where infrastructure development does not take account of special requirements of the visually impaired. This leads to visually impaired people being dependent on others. The need is especially huge for new environments and traffic places, and in comprehending road signs like 'STOP' or traffic signals. These reasons have led to high poverty and unemployment, resulting in poor quality of life and adding to the misery of the individuals suffering from blindness.

BY VISHESH K. **PANJABI** AND DR. BHARAT **LOHANI**

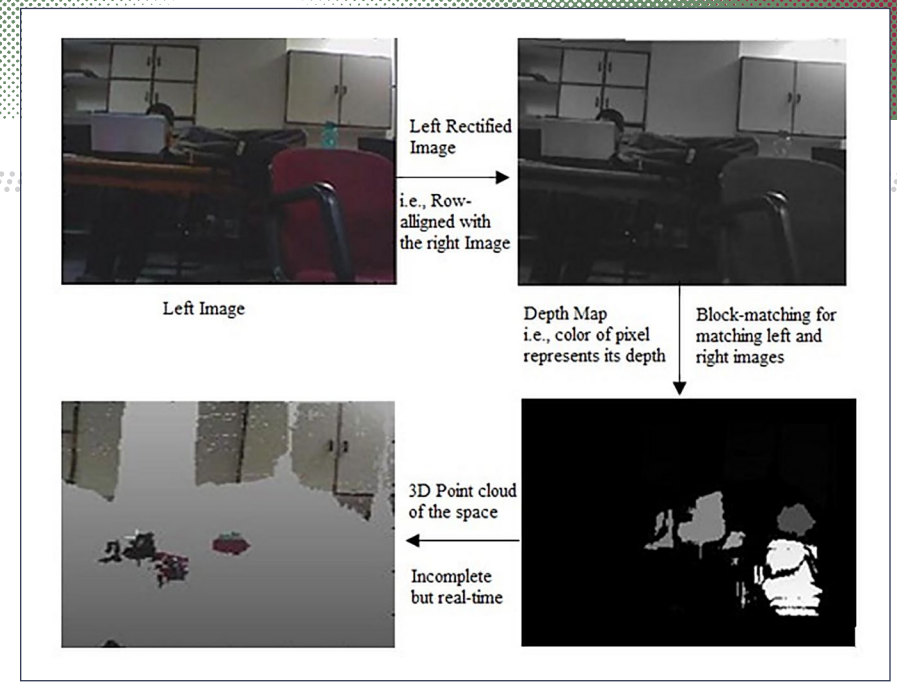

Figure 2: Generating a 3D point cloud from left-right image pair.

The most widely used device by the visually impaired is a 'cane' (stick), which has a field of view equivalent to a point, making the blind person vulnerable to any obstacle lying at some height above the ground. Apart from this, one hand of the person is always busy in carrying the cane which makes it an irritating experience.

However, with advances in computer vision and audio systems, this dependence on the cane can be eliminated. In this work we are trying to make a visually impaired person completely autonomous as far as local locomotion in a new environment is concerned.

Technical Description

The entire methodology of creating the obstacle avoidance system consists of:

1. Using stereo pairs acquired from a web-cam system and implementing stereo-vision algorithms to generate a 3D point cloud.
2. Using the generated 3D point cloud to identify and characterise obstacles lying in the path of the user.
3. Converting and conveying this processed output in the form of an audio beep, thus minimizing information overloading.

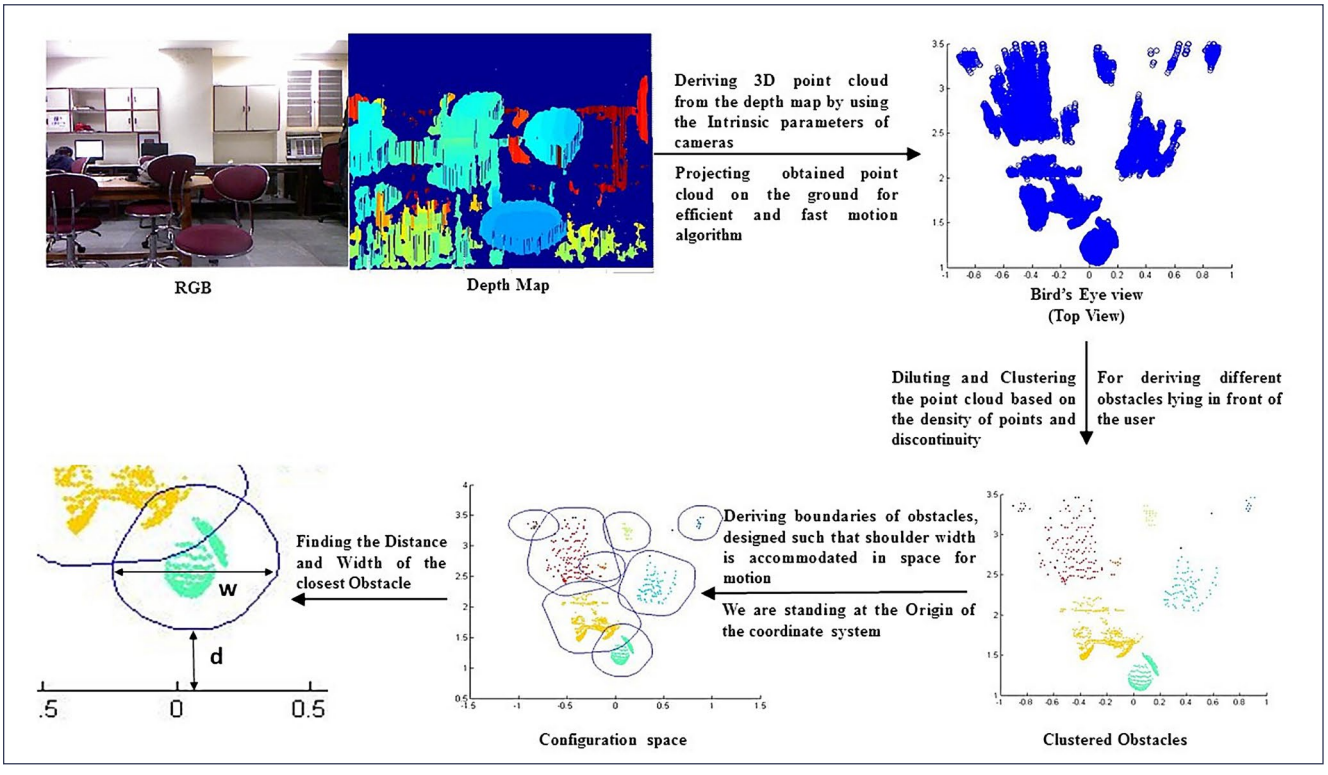

Figure 3: Steps to derive obstacle characteristics from the point cloud.

Creating 3D Point Cloud From the Stereo-Pair

Since real-time computations are a must for this application, we have used a pre-calibrated stereo-system instead of flexible (where the cameras are not fixed) systems. To ensure rigidity and light weight requirements, acrylic was used to make the outer case of the system, in which two-web cams were placed to acquire the real-time stereo-pairs as shown in (**Figure 1** left-bottom). The baseline of this stereo-rig was 80mm in order to set the widest possible baseline that can be easily worn by a person as a normal spectacle.

After this, the EO (exterior orientation) and IO (interior orientation) calibration of the stereo-rig and rectification mapping calculations were carried out, making the system ready for real-time generation of the 3D point cloud. This was followed by implementation of a block-matching-based correspondence algorithm on the greyscale stereo-pair in order to derive the 3D coordinates of each pixel in the image. **Figure 2** provides a figurative summary of the entire procedure.

Processing Point Cloud to Characterise Obstacles

Firstly, we eliminate the points which lie under the feet and above the head of a user. The remaining points are then projected onto the ground. This converts the 3D view into a top-view of the scene, where the user is standing at the origin. After this step we differentiate between each individual obstacle through a density-based clustering algorithm. The individual clusters identified in this step are essentially the obstacles. The boundary of each individual cluster (obstacle) is expanded by a half-shoulder width (configuration space) to account for the shoulder width of a person. Thus, henceforth, a person can be considered as a dot of zero size and if a path can be found where the dot can pass through successfully (without venturing into the configuration space) a person can also pass through that path. The closest obstacle is the one which is the closest to the person and needs the first attention. So we isolate the closest obstacle and derive its distance and width. **Figure 3** describes the same procedure, figuratively.

Figure 4: Corridor with obstacles used for experimentation.

Conveying the Distance and Width through Auditory Channel

After this, the distance to the closest obstacle is represented in the form of amplitude '*A*' of an audio beep, given by the following formula, where the constant comes from our experimentation and literature in acoustic engineering.

$$A(d) = 0.15/d$$

Testing the System

Given the critical nature of the application, it was important to experimentally evaluate the effectiveness of the developed system. In order to keep the path taken consistent among all the participants, we chose the corridor as shown in **Figure 4** and placed chairs as obstacles for the participating individuals.

The participants were instructed to travel from one end to the other, blindfolded using the developed system. The goal was to avoid bumping into any of the walls or obstacles. Before the test was conducted, the participants were told how the system worked and also were given a test ride with the system until they felt acclimatised to the device. As currently the vertical field of view of our device is limited we made the blind-folded person move on a chair with wheels, as shown in **Figure 1** right.

Out of the seven obstacles in the scene the participants on an average collided with one. Moreover, the experience sharing revealed that since the participants were blindfolded temporarily and not blind by birth, this affected their heuristic to use the system effectively, especially in the beginning. However, as one of the participants claimed "once the heuristic to use this system becomes clear that is when you realise how to navigate with it".

This experiment, proved the potential of our system as a successful replacement for the cane.

Frugality of the System

The prototype is prepared using two web-cameras, a pair of earphones and a processor (laptop). The final product will consist of two web-cameras, a pair of earphones and an FPGA (Field-Programmable Gate Array) implementation of the algorithms. This results in a system cost of less than USD 80, which could be further reduced with higher volume manufacturing.

Dr. Bharat Lohani is a Professor at IIT Kanpur and focuses his research in LiDAR technology development and applications.

Vishesh K. Panjabi is a masters student at IIT Kanpur, his research focusses on 3D Computer Vision and its Applications in Geomatics.