A Portable Driver Assistance System Headset Using Augmented Reality

Shantanu Misra[1*], Vedika Parvez[2], Tarush Singh[3], E. Chitra[4]

[1]*Dept. Of Electronics & Communication, SRM Institute Of Science And Technology, Kattankulathur, Tamil Nadu, India.*
[2]*Dept. Of Electronics & Communication, SRM Institute Of Science And Technology, Kattankulathur, Tamil Nadu, India.*
[3]*Dept. Of Electronics & Communication, SRM Institute Of Science And Technology, Kattankulathur, Tamil Nadu, India.*
[4]*Dept. Of Electronics & Communication, SRM Institute Of Science And Technology, Kattankulathur, Tamil Nadu, India*
Corresponding Author E-Mail:Shantanu_Misra@Srmuniv.Edu.In

Abstract

Vehicle collision leading to life threatening accidents is a common problem which is incrementing noticeably. This necessitated the need for Driver Assistance Systems (DAS) which helps drivers sense nearby obstacles and drive safely. However, it's inefficiency in unfavorable weather conditions, overcrowded roads, and low signal penetration rates in India posed many challenges during it's implementation. In this paper, we present a portable Driver Assistance System that uses augmented reality for it's working. The headset model comprises of five systems working in conjugation in order to assist the driver. The pedestrian detection module, along with the driver alert system serves to assist the driver in focusing his attention to obstacles in his line of sight. Whereas, the speech recognition, gesture recognition and GPS navigation modules together prevent the driver from getting distracted while driving. In the process of serving these two root causes of accidents, a cost effective, portable and holistic driver assistance system has been developed.

Keyword: Augmented reality, driver assistance system.

1. Introduction

The Accident report 2016 [1] states that as many as 17 people die in 55 road accidents per hour on average each year in India. Though overall road accidents declined leading upto 2016, the fatalities increased by 3.2 percent, meaning more than 400 people lost their lives daily. 1.5 lac people lost their lives in 2016 due to half a million road accidents that took place in India. A survey states that in India, most people are killed in road accidents rather than by diseases.

In accordance with such statistics recorded each year, it is imperative that action be undertaken for road safety measurement. One such attempt was successful through the obstacle detection system, which was designed to reduce rates of occurrence of accidents. This was earlier achieved by introducing new and efficient technologies like radar, lasers, cameras and ultrasonic sensors to detect obstacles in front or behind the vehicle. However, due to the lack of feasibility and high cost of such technologies, it wasn't an acceptable solution for the pressing situation at hand. The work however, had been proposed after analyzing the Indian road conditions[2] and densely populated areas which posed to be the main factors contributing to this approach. Driver distraction and inattention are evident causes of automotive collisions. New sensing approaches are required to enable driver-assistance systems to address these problems. Need of the hour is for a cost effective, portable and holistic driver assistance system in order to combat these issues and resume road safety measures in India.

A review of the different technologies used for obstacle detection is given in Section 2 of this paper. Section 3 deals with the overall headset design. Methodologies of the five modules is given in Section 4. Section 5 shows the results obtained from the different modules.

One of the most complicated tasks in real time systems is obstacle detection in automobiles. In 1988, the Delco System Operations paved the way for obstacle detection in which the obstacles on the road were sensed. Followed by the Delco is the Santa Barbara Research Corporation, California which also set forth better methods in detection. Many advanced driving assistance systems are being proposed in the past years to ensure safety[3].

New obstacle detection systems are familiarized with sensors like wave radar, Light Detection and Ranging (LIDAR) sensor, and ultrasonic sensor. Analysis like tracking of objects from images, moving away from obstacle and automatic measurement of the obstacle from the vehicle are done[4]. Latest technologies like augmented reality (Augmented Reality (AR)), image processing, gesture recognition and speech recognition are used in this system to detect obstacles around the vehicle. In order to combat inattention of the driver, one of the main parts of this invention is to analyze any kind of obstacle on the road. For example, stationary objects, people moving around are certain things that get in the passage way which are also considered as obstacles. These obstacles are detected by the system and alert the driver through a visual detection window for safe and easy navigation. Presence of pedestrians is detected by object detection method. Secondly, distraction of the driver, owing to the use of mobile phones, etc. is combated vide the incubation of a Global Positioning System (GPS) navigation system, gesture recognition module and a speech recognition module in the invention.

2. Headset Description

The headset consists of an Organic Light Emitting Diode (OLED) module which acts as the display unit for the headset. The prototype eyepiece is made from plexiglass and has a cuboidal structure. The plexiglass piece is inclined at an angle of 45° to the OLED module attached. Due to its inclination, internal reflections take place enabling the headset to display output on the plexiglass. The heart of the system is a 64-bit microprocessor which powers all the modules functioning in the system. The Global Positioning System GPS, OLED, magnetometer modules and a microphone are attached to the processor that function based on the external stimuli received from the environment. The processor also powers a 32-bit microcontroller which is used to facilitate gesture recognition. Communication between the microcontroller and microprocessor takes place over Universal Synchronous and Asynchronous Receiver-Transmitter Universal Synch/Asynch. Receiver Transmitter (USART) communication protocol using values from the Inertial Measurement Unit (IMU) and flex sensors.

Upon powering the headset vide a lithium polymer battery, the pedestrian detection module ensures that the driver detects pedestrians and is adequately aware of the obstacles in his path. In a situation where the driver is unable to respond to the stimuli emanating from the obstacle the driver alert system is activated. This module triggers a warning signal for the user through a portable speaker. Apart from first two system, a speech recognition and gesture recognition system is also activated which sets up an interface between the user and the headset allowing the user to cater to peripheral functionalities like responding to text messages, calls, etc. The speech-to-text module also enables the user to operate the GPS module and set a destination by relevant providing speech stimuli. The directions for the destination are shown on the heads up display in line with the driver's vision to eliminate distraction and reduce the risk of encountering road accidents.
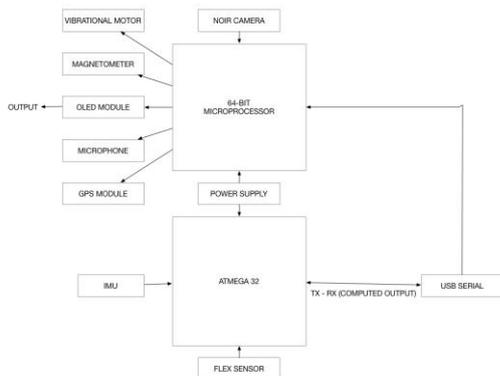

Figure 1: Headset block diagram

3. Module Implementation

Pedestrian Detection Module

As shown in Fig. 2, we use a No Infrared (NoIR) camera to capture raw images. The Raspberry Pi and camera are connected to each other using a Camera Serial Interface (CSI) which carries pixel data.[[8]. The raw image captured by the NoIR camera is converted into grayscale value from the Red Green Blue (RGB) values of the video feed. Each frame of the image is processed by HOG (Histogram Oriented Gradient) algorithm which maps the gradient values used to map the edges of the obstacle[9]. HOG is opted as descriptor as it outperforms existing edge based and gradient based descriptors in terms of speed, ease of implementation and optimization during training[10].

HOG scans the image in a sliding window manner. The smaller the window, larger the dataset made. The returned dataset is such that it assigns '1' for an edge value while assigning '0' for the rest. HOG works along with Support Vector Machines to classify the

images captured by the NoIR camera. It is worth mentioning that using Support Vector Machines (SVM) in combination with HOG provides a very robust method of detecting humans[11][12] the duo of which has been leveraged upon in this innovation. The edge coordinates are used to mark obstacles, highlighting them on the OLED screen by forming a box around it. The three parameters used to detect the human obstacle are its - padding, scale and winstride. As the system uses SVM, it can map people of any size and shape in real time[13].
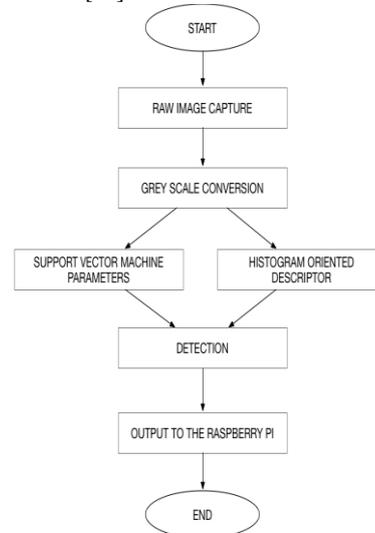

Figure 2: Pedestrian Detection Module

Speech Recognition Module

As seen in Fig. 3, the speech recognition process is initiated the moment speech input is given by user through the microphone. The audio to be transmitted is first encoded with the Free Lossless Audio Codec (FLAC) encoder, which enables transmission of audio file that have been compressed without major losses[14][15]. The encoded file is sent to Google API where using audio analysis and machine learning algorithms - Long Short Term Memory (LSTM) and Neural Machine Translate (NMT), it deciphers the text in audio file and converts it to text format. One key advantage of Google Application Program Interface (API) is its low 'word error rate' when compared to other speech recognition systems[16][17].
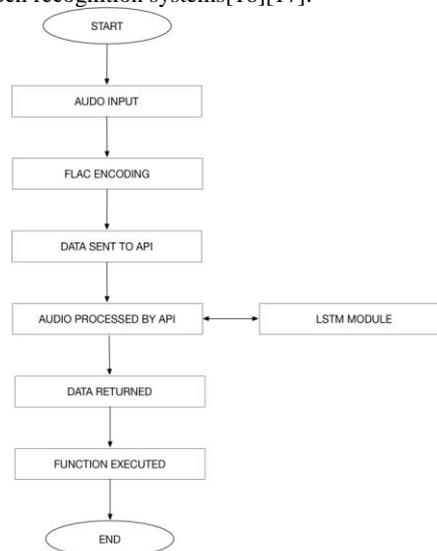
Figure 3: Speech Recognition Module

Upon obtaining the input audio in a text format, the API checks the language in which the audio is, and the language in which the output is expected. Using LSTM specifically over other sequence prediction systems is worth mentioning because LSTM removes complex time lags, stores data in its memory for a longer time and

can also be trained quickly. Additionally, LSTM gives a low test error value when compared to other available systems [18][19][20]. Once all the words are successfully matched from the received alternatives, a sentence is formed. Following this, the system deciphers which sentence is more similar to the audio file received. After that, the system returns the output in the form of a string. Once reception of the string is successful, the program matches the string to the provided list of predefined functions and the function being called is executed.

GPS Navigation Module

The input for the GPS navigation system is obtained from the speech recognition module. GPS modules use navigation for assisting drivers while on the road[21]. The GPS module sets the end position by taking in input from user through a prespecified speech format. By default the system uses the current location as a start position. To achieve this, the GPS module gives a request signal for co-ordinates via the RX pin of the GPIO and finds out the longitude and latitude cellular communication. The values are sent back to the program via the TX pin. The received coordinates are used as the starting point location. The end point variable is filtered from the speech-to-text string. The two variables are then passed to maps application which initiates navigation. The information such as distance to be travelled, direction to turn towards is displayed on the heads-up display of the driver. A magnetometer acts as the compass to determine user direction.

Driver Alert System

As exhibited in the figure, the driver alert system is a surveillance system which detects rate of approach of obstacles and triggers an alarm based on values received. The following steps are involved in the working of this module:

1. The system obtains coordinates for the obstacles from the pedestrian detection module. On receiving them, the alert system calculates the area of the box and stores the value in an array of floating points.
2. If the value obtained is positive, the obstacle is assumed to be going away from the user. If the value obtained is negative, the obstacle is coming towards the user.
3. Based on rate of change of the values, the speed of approach is calculated[22]. These values are compared to that of the average breaking system of a moving vehicle given by the equations -

$$Lo = LA + LB \quad (1)$$

$$LA = A \times V2 \quad (2)$$

$$LB = B \times V \quad (3)$$

Where Lo signifies the distance at which vehicle stops after breaking, LA is breaking caused by tyre-road contact. LB signifies the breaking due to the breaking constant B. A is the coefficient of friction, V is the velocity of the vehicle and V2 is the velocity of the vehicle after breaking.

4. If the obtained values are within the preset range, the alarm doesn't trigger. However, if the obtained values are out of the range, the alarm is triggered.
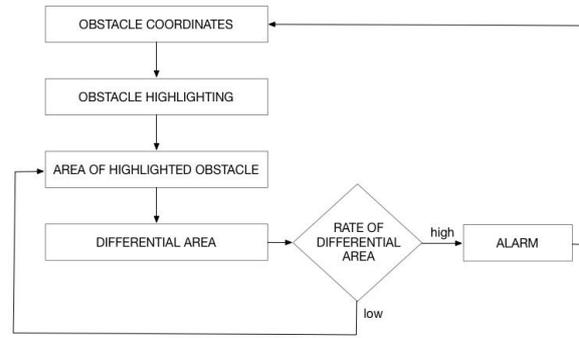

Figure 4: Driver Alert System

Gesture Recognition Module

The Inertial Measurement Unit working in tandem with an arduino provides a way for the driver to use gestures to control functions of a mobile device[23][24]. The IMU uses i2c protocol which has one data pin and one clock pin. Based on the clock cycles, the inbuilt multiplexer IC cycles between the received values of the gyrometer and accelerometer values. After receiving the values, the MUX appends all the values and sends the values using the data pin. The flex sensor acts as a switch such that when it gives the values of one, only then the arduino will transmit the values based on which the python interpreter will decide the action to be performed. The arduino transmits the values using the USART communication protocol at a baud rate of 9600. To make the data transmission seamless and without any data loss, a time delay of 500 milliseconds is given. This ensures that correct values are received by the python interpreter for further output.
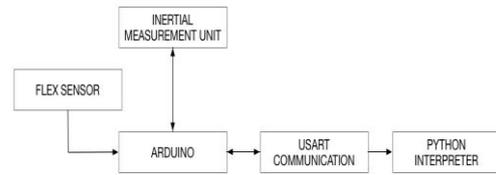

Figure 5: Gesture Recognition Module

4. Result

All the modules were tested individually and in unison. The results were cross tabbed and are shown as follows.
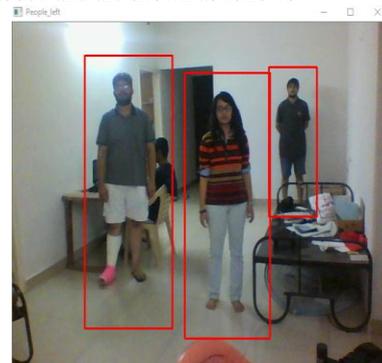
Figure 6: Pedestrian detection output

Table 1: Parameters for Pedestrian Detection

Parameter	Value
Maximum distance for detection	100 metres
OLED - Frames per second count	28
WinStride	4,4
Scaling Factor	1.14
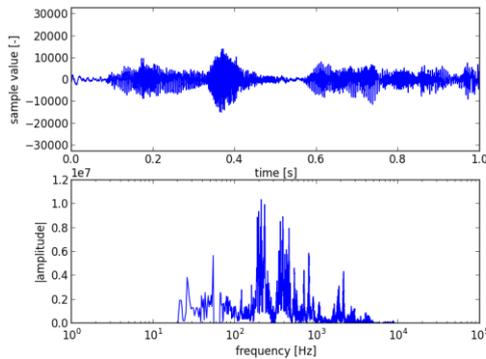
Figure 7: Audio Waveform - Speech Recognition Module

**Figure 8:** GPS module - SRM IST to whitefield, Bangalore
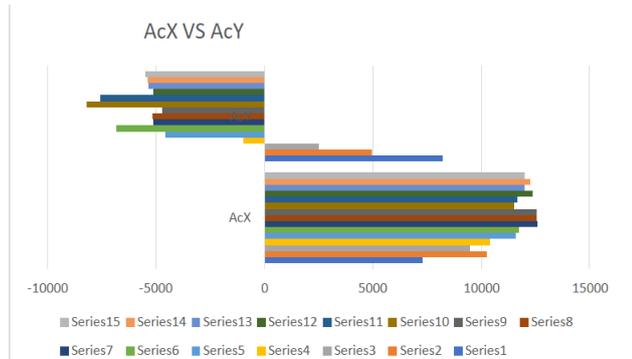
Figure 9: Driver hand movement in XY plane - Gesture Recognition Module

The model undertaken in designing of the driver assistance system headset established that the working of the in-sync modules - pedestrian detection, driver alert system, speech recognition, gesture recognition and GPS navigation made for a portable driver assistance headset, powered by augmented reality. Its portability, ability to work under unfavourable weather conditions and ability to work in any road conditions made it a holistic solution to the pressing issue of road accidents at hand. Additionally, since most of the components used were cost-effective, it also made the model economically viable.

Acknowledgment

References

[1] Mohan D, Road accidents in India", *IATSS research*, Vol.33, No.1, (2009), pp.75-79.

[2] Dabbour E & Said E, "Proposed collision warning system for right-turning vehicles at two-way stop-controlled rural intersections", *Transportation research part C: emerging technologies*, Vol.42, (2014), pp.121-131.

[3] Ujjainiya L & Kalyan Chakravarthi M, "Raspberry Pi Based Cost Effective Vehicle Collision Avoidance System Using Image Processing", *ARPN J. Eng. Appl. Sci.*, Vol.10, (2015), pp.3001-3005.

[4] Karoui I, Isabelle Q & Michel L, "Automatic sea-surface obstacle detection and tracking in forward-looking sonar image sequences", *IEEE Transactions on Geoscience and Remote Sensing*, Vol.53, No.8, (2015), pp.4661-4669.

[5] Datcu D, Marina C, Stephan L, David MO & Mikael W, "Virtual co-location to support remote assistance for inflight maintenance in ground training for space missions", *15th International Conference on Computer Systems and Technologies*, (2014), pp.134-141.

[6] Lukosch S, Heide L, Dragoş D & Marina C, "Providing information on the spot: Using augmented reality for situational awareness in the security domain", *Computer Supported Cooperative Work (CSCW)*, Vol.24, No.6, (2015), pp.613-664.

[7] Azuma RT, "A survey of augmented reality", *Presence: Teleoperators & Virtual Environments*, Vol.6, No.4 (1997), pp.355-385.

[8] Back A & Deady C, "Introducing the Pi NoIR Camera", *The MagPi*, Vol.18, (2013), pp.16-18.

[9] Dalal N & Bill T, "Histograms of oriented gradients for human detection", *IEEE Computer Society Conference on Computer Vision and Pattern Recognition,* Vol.1, (2005), pp.886-893.

[10] Zhu Q, Mei CY, Kwang TC & Shai A, "Fast human detection using a cascade of histograms of oriented gradients", *IEEE Computer Society Conference on Computer Vision and Pattern Recognition,* Vol.2, (2006), pp.1491-1498.

[11] Yadav RP, Vinuchackravarthy S, Krishnan K & Sunita PU, "Implementation of robust HOG-SVM based pedestrian classification", *International Journal of Computer Applications*, Vol.114, No.19, (2015).

[12] Bristow H & Simon L, "Why do linear SVMs trained on HOG features perform so well?", *arXiv preprint arXiv:1406.2419*, (2014).

[13] Chen Z, Kai C & James C, "Vehicle and Pedestrian Detection Using Support Vector Machine and Histogram of Oriented Gradients Features", *International Conference on Computer Sciences and Applications (CSA),* (2013), pp.365-368.

[14] Muin FA, Teddy SG, Mira K & Elsheikh MAE, "A review of lossless audio compression standards and algorithms", *AIP Conference Proceedings*, Vol.1883, No.1, (2017).

[15] BRivero C & Prabhat M, "Lossles audio compression: A case study", Technical Report 08-415, Department of computer and information Science and Engineering, University of Florida, Gainesville, FL32611, USA, (2008).

[16] Këpuska V & Gamal B, "Comparing speech recognition systems (Microsoft API, Google API and CMU Sphinx)", *Int. J. Eng. Res. Appl.*, Vol.7, (2017), pp.20-24.

[17] Kulkarni DH, "Review on Recent Speech Recognition Techniques", *International Journal of Scientific and Research Publications*, Vol.3, No.7, (2013).

[18] Sak H, Andrew S & Françoise B, "Long short-term memory recurrent neural network architectures for large scale acoustic modeling", *Fifteenth annual conference of the international speech communication association*, (2014).

[19] Graves A, Abdel RM & Geoffrey H, "Speech recognition with deep recurrent neural networks", *IEEE international conference on Acoustics, speech and signal processing (icassp),* (2013), pp.6645-6649.

[20] Hochreiter S & Jürgen S, "Long short-term memory", *Neural computation*, Vol.9, No.8, (1997), pp.1735-1780.

[21] Gallet A, Spigai M & Hamidi M, "Use of vehicle navigation in driver assistance systems", *Intelligent Vehicles Symposium,* (2000), pp. 492-497.

[22] Iizuka H & Kunihiko K, "Automatic braking system with proximity detection to a preceding vehicle", U.S. Patent 5,278,764, (1994).

[23] Benbasat AY & Joseph AP, "An inertial measurement framework for gesture recognition and applications", *International Gesture Workshop*, Springer, Berlin, Heidelberg, (2001), pp.9-20.

[24] Mantyla VM, Mantyjarvi J, Seppanen T & Esa TH, "and gesture recognition of a mobile device user", *IEEE International Conference on Multimedia and Expo,* Vol.1, (2000). pp.281-284.