Integration of Commercial GPS with On-Board Vision Module for the Navigation of Unmanned Vehicles in a Variable Intensity Domain

Sudhir R Posireddy, Sabri Tosunoglu
Florida International University
Miami, Florida, USA

ABSTRACT

Today, many researchers address the challenging problems of robotics such as visual recognition, navigation and machine learning. Over the past few years the competitors of DARPA Urban Challenge have demonstrated significant advancements in these areas and the iRobot corporation has become the key leader in building commercial government and industrial Unmanned Ground Vehicles (UGV’s) for various military applications. Over the past two decades, this company has maintained their unique position with their low-cost compact unmanned mobile robots (UGV’s). The goal of this paper is to design an advanced unmanned mobile robot using low cost modules to closely challenge today’s high cost systems used for developing UGV’s. This goal is achieved by seamlessly interfacing low cost vision module [1] and commercial GPS receivers with robot’s brain. Since commercial GPS has a voice command output, in this paper we develop an innovative technique of interfacing low cost GPS receivers with HM2007 speech recognition IC, which enables us to precisely recognize the direction in which the robot has to travel and the position at which it has to stop with a standard speaker-dependent voice recognition system.

Keywords

1. INTRODUCTION

Among the many open issues to be addressed in the development of the mobile robot navigation system, path planning is one of the major issues. A great number of different techniques have been developed in order to carry out efficient robot path planning [19, 20, 21]. The iRobot corporation has developed many different kinds of unmanned mobile robots for different applications varying from service industry to military robots.

Path planning for robots and manipulators is a research area where new contributions have been reported by researchers for decades. Over the years, several groups of methods dealing with the path planning problem have been developed. A short evolution of the state of the art can be found in [2], where Latombe lists the main landmarks on path planning until 1990 and in the field of obstacle avoidance, there are several Ph.D. dissertations [4, 5, 6] and reports [7] where exhaustive classification and description of the different methods have been carried out.

Mobile robot systems use sophisticated equipments for successful execution of the planned path, e.g., the mobile robot system at CMU NAVLAB [8] consists of several modules, namely color vision, to see the road; 3-D vision, looking for obstacles; the pilot, predicting road locations for color vision and planning vehicle trajectories; the Helm trajectory execution manager; the controller, which handled the real-time tasks of vehicle driving; and the map navigator, responsible for route planning. Similarly, in the work carried out in Jet Propulsion Laboratory [9], for obstacle detection, the mobile robot system consist of a forward-looking stereo camera, infrared and sonar proximity sensors looking in various directions, and the scanning laser rangefinder. To provide adequate coverage, the stereo cameras have a field of view of 97×74. The stereo imagery is processed into 80×60 pixel disparity maps on the vision stack using algorithms developed previously at JPL [10]. To provide adequate dynamic range and control of image exposure for operation from bright sunlight to dim indoor lighting, cameras are required to have software-controllable exposure time.

But as pointed out by the authors of the above paper [8], most existing road trackers work well for only a particular road, or only under benign illumination. In the presence of disturbances such as disappearing features or change in illumination the models do not stand up to the test. In order to address such issues, the researchers in this work [8] have proposed to build and use explicit road models. In this work [8], specific models are made to track road edge markings (white stripes), road center lines (yellow stripes), guard rails, shoulders, type and color of road surface, location of vehicle on road, location of stripes etc. In addition, they have also proposed to have explicit models for shadows, local changes in road surface, e.g. patches, global illumination changes, such as the sun going behind a cloud, etc.

Similarly, the researchers at JPL have built specific algorithms to deal with shadows [9]. In this work, to cope with strong shadows and significant appearance variations, the edge detection algorithm takes a “least commitment” approach to finding the near-parallel straight lines that are the edges of the steps. First, a Canny edge detector is run with a low gradient magnitude threshold to find even weak edges; edges are then linked into straight line segments. The dominant orientation of all of the edges is found by histograming the edges in all of the edges and choosing the greatest peak within 45 degree of horizontal; all edges with orientations further than some threshold from the dominant orientation are then discarded. Since some steps may be detected as multiple short line segments, the remaining edges are
filtered to merge nearly collinear segments. Finally, any edge that is still less than 1/4 of the image width in length after merging is discarded. Some distracting edges with inconsistent orientations can still remain at this stage.

Similarly, the researchers have presented specific algorithms for illumination invariance among a few other things for improved localization on legged robots [11]. As pointed out by the authors, color is often one of the most informative visual cues in the environment and color segmentation is an inherently uncertain operation. Furthermore, under changing illumination conditions, the same pixel values may represent different colors. The authors presented algorithms for autonomous color calibration and a method aimed at directly achieving illumination invariance.

Additionally, the author has presented SIFT algorithm to make the path planning algorithm invariant to change in illumination among a few other parameters for robot localization [12].

This discussion above concludes that varying illumination in the domain of the mobile robot plays an important role on its path following and as suggested by many authors this illumination and intensity problems of light can be gained over by adapting the high cost algorithms which are not feasible for developing low-cost unmanned vehicles.

The authors have previously developed a revolutionary low-cost technique to draw the optimum path to the selected destination by integrating the potential field concept [13] and low-cost on board vision module [1]. Though we successfully discussed and demonstrated the cost effective solutions for developing a mobile unmanned robot [1], we still have a challenging problem in dealing with the position accuracy of standard digital GPS modules.

In this study, we discuss a reliable cost effective technique to resolve the issues of position accuracy with standard digital GPS module by integrating commercial GPS (such as Tom-Tom, Gramin, etc.) with on-board vision module equipped with voice recognition system. This enables us to take the direction information from the voice commands of commercial GPS system and the obstacle information from the vision module. This information can be used with path planning algorithm [1] to determine the optimum path to the destination by avoiding all the obstacles.

The 3-D solid model of the proposed robot is shown and the algorithm is implemented in real time to test the suitability of the robot for a real environment.

The resulting system makes it possible to utilize our two important independent modules: (1) On board vision system, (2) Commercial GPS module. In this work, these modules are integrated together with low end microcontrollers to implement path planning algorithm for GPS based vision guided navigation.

2. COMMUNICATION BETWEEN MODULES

This is the crucial part of the system which is as important as arteries and veins of our body. The communication and synchronization of data between the modules plays the major role for getting the system work as expected. Block diagram in Figure 1, explains the working of our “VizzBot-I”. Bayer data from camera is sent to AT Mega 8 and is future processed for pixel information. ATTiny 12 is used to synchronizes the communication between Camera (OV 6620) and AT Mega 8. This synchronization is very important to avoid the loss of any pixel information when grabbing full image from the camera.

Commercial GPS Module from Tom-Tom shown in Figure 2, is used to guide the robot from one location to another location. Instead of opting for standard parallax GPS system, we demonstrated the high level cost effective solution with commercial GPS unit, by integrating it with voice recognition system. The commercial GPS will speak out the respective commands whenever we need to change the direction of travel. These commands are recognized by voice recognition unit and respective information is send to P89V51RD2 for future processing to command the actuators, so as to move the robot in required direction. The height of synchronization for this voice recognition unit is less when compared with the synchronization level required for vision module. At any stage of program execution the obstacle data from vision module is given highest priority when there is parallel information from both the modules to process the data.

<table>
<thead>
<tr>
<th>Decoded Direction of Travel Information:</th>
</tr>
</thead>
<tbody>
<tr>
<td>000 - Straight Condition</td>
</tr>
<tr>
<td>001 - Left Condition</td>
</tr>
</tbody>
</table>
The designed vision system [1] provides high-level obstacle information extracted from the camera image to an external processor P89V51RD2 that controls a mobile robot to avoid the obstacles. P89V51RD2 the brain of VizzBot-I, uses the obstacle information from vision module and direction information from GPS to process the path planning algorithm. This algorithm delivers the optimum path in which the robot has to travel to reach the destination by activating the motors accordingly.

From above discussion it is understood that there are two modules in developing this unmanned robot: (1) Image processing block for obstacle information, and (2) GPS for direction information. Let us first start our discussion with image processing module.

3. ON-BOARD IMAGE PROCESSING

As discussed in [1] we modified the algorithm in AVRcam to best meet our required goals. The AVRcam, shown in Figure 3, is a real-time on-board image processing engine that utilizes an Omnivision OV6620 [1] CMOS image sensor, mated with an Atmel AVR mega8 microcontroller to perform all of the image processing tasks.

One of the key issues in image processing is the effect of changing light conditions in the environment. This tends to complex the algorithm even for simpler applications and implementing image processing algorithms on a low resolution image with varying light conditions makes it very tedious to develop a real time adaptive system.

We propose a novel method of dealing with varying light conditions by using YUV color space so as to win over the major issues of image processing. This enabled us to develop an easily adaptive system with high repeatability and feasibility.

3.1 Advantages of YUV Color Space

Raw Bayer data from camera module is send to AT Mega 8 and future processing of pixel information is done in YUV color space. Usage of Intensity- Value (Y-Luminance, UV – Chrominance Value) color space adds two advantages to our model: (1) to overcome the effect of varying light conditions, (2) reduce the computation cost of processing, by reducing the no of bits to be processed per each pixel.

To process an image in RGB color space, shown in Figure 4, we need to process all the 8 bits of each pixel of each ‘R’ ‘G’ ‘B’ planes, so in total 24 bits of data has to be processed for one pixel information. This is huge amount of data to be processed for an image of size 176x144, i.e. 176x144x24 (608256) bits of data.

For converting RGB color space to YUV space we use the following mathematical transformations:

\[
Y = W_R \times R + W_G \times G + W_B \times B
\]

\[
U = 0.436 \times \frac{(B - Y)}{1 - W_B}
\]

\[
V = 0.615 \times \frac{(R - Y)}{1 - W_R}
\]

Where \(W_R\), \(W_B\) and \(W_G\) are the weights whose values are:

\[
W_R = 0.299
\]

\[
W_B = 0.114
\]

\[
W_G = 1 - W_R - W_B = 0.587
\]

Processing image in YUV color space stores all the luminance values in Y plane and chrominance value information in UV planes. So, just by not considering the data from Y plane we can eliminate the effect of lightning on the environment and since we are not considering the data in ‘Y’ Plane, total no of bits of data to be processed for one pixel information will be reduced by 8 times. This helps us in saving lots of processing power.

3.2 Image Processing Algorithm in AVRCam

The algorithm in Figure 6 is implemented continuously for each and every captured image till we reach the destination. The technical details and working of AVRcam is discussed elaborately in [1].
Eliminate Data in ‘Y’ Plane
Set Registers for AVR Cam to operate in YUV Color space
Get Raw Bayer ‘UV’ image data
Process for pixel information

Figure 6. Algorithm Implemented in On Board Vision Module (AVR Cam)

Running along this Algorithm will get us an image as shown in Figure 7, which is free from all luminance values (‘Y=0’). It means that whatever be the lighting conditions the pixel values in Figure 7 will be almost all constant. This way we can avoid implementing complex algorithms for simpler and adaptive applications.

Figure 7. Image with No Luminance Values

4. INTERFACING COMMERCIAL GPS FOR DIRECTION INFORMATION

In this section we discuss the reliable cost effective technique to resolve the issues of position accuracy in [1] by integrating commercial TomTom GPS with on-board voice recognition system (HM 2007). This enables us to take the direction information from the voice commands of commercial GPS system. This information can be used as an input to P89V51RD2, the brain of Vizzbot-I for deciding the direction in which the robot has to travel.

The HM 2007 developed by HUALON Microelectronics (HM) Corporation is the most popular and commercial single chip CMOS voice recognition Large Scale Integration (LSI) circuit with an on-chip analog front end, voice analysis, recognition process and a system control function. The operation modes provided by the HM2007 are manual mode and CPU control mode. The CPU mode of operation is used to interface the HM2007 to a programmable controller.

The HM2007 can recognize either 40 0.9-second long words or 20 1.92-second long words or phrases. The recognition technology used by this IC is of the speaker dependent discrete type meaning that it can only recognize words spoken by a single person and the speaker must "train" the system with each word to be recognized. Because GPS always delivers a standard voice and fixed commands we opted to use this IC in our design to recognize the direction commands from GPS. The pin configuration for the HM2007P IC is displayed in Figure 8.

The HM2007 speech recognition processor in SR-06/07 toolkit is shown in Figure 9.

4.1 Training and Recognition

We must first train the speech recognition processor of the SR-06/07 toolkit to be able to recognize any given input voice commands. The toolkit is equipped with a keypad, a microphone and two seven segment LED display. The microphone is provided for the analog input voice command into the speech recognition processor. The speech recognition processor stores templates of the speaker’s voice characteristics such as tone, inflection, and accent. For speaker dependent type systems the same person that trained the system will have to speak again during the testing phase in order for their voice command to be recognized. The voice command is recognized by matching the characteristics of
the spoken word to those already stored in the template. The output of the speech recognition processor consists of two 4 bit BCD numbers that are displayed on the seven segment LED displays. However this output is also available on the 10 pin header which uses eight pins for the two 4 bit BCD and one pin for VCC and the other pin for ground. These two 4-bit BCD numbers will then be transferred to the microcontroller which will then decode the acquired data as discussed above and transmit the corresponding bytes of information L293D the motor controller IC to drive the motors accordingly.

We further increased the accuracy of the voice system by filtering the noise out from the voice commands of GPS. We first input the commands from GPS to the microphone which is then passed through the Active filter circuit with class D amplifier’s output to speaker. The output of this filter is then spelled out from the loud speaker as an input to the microphone of the voice recognition circuit. This way, we first process the voice command for noise removal and then train the speaker dependent system. And we also trained the most commonly used commands in more memory slot compared to the commands that are sparingly used so that we could have more hit rate and reliability of the system.

![Figure 10: Active Filter Circuit with a Class D Amplifier Circuit](image)

This circuit uses MAX-9727 quad-audio-line driver, IC1 which combines separate single-ended filters—one for each of the BTL outputs’ phases—with a third amplifier that provides a difference signal with additional filtering. The first stage of each single-ended-filter section contributes the complex-conjugate pole pair of a third-order, 30-kHz multiple-feedback Butterworth filter, for which many design guidelines and equations are available. Each third-order-filter section comprises a complex-conjugate pole-zero pair and one real pole. To improve the match between the signal paths, the two separate multiple-feedback filters share a real pole, which 470-pF capacitor C1 and 11-kΩ resistors R1 and R6. The circuit implements that pole as a difference amplifier, thereby producing a filtered output that presents a single-ended version of the BTL amplifier’s outputs. Auxiliary components in this circuit can be selected according to remove the noise from the voice commands effectively. This processed data can now be used for future processing the direction information which results in more reliable output.

5. IMPLEMENTATION IN A MOBILE ROBOT

Figure 11(a) show the 3-D solid model of the mobile robot designed to test this algorithm. Two metallic bases of rectangular cross-sections are provided for mounting the on-board camera and the circuit board. A castor wheel, in the front is mounted to provide free motion to the robot. The motors along with the clips to hold the same are shown to drive the wheels. Figure 11(b) shows the first version of VizzBot which uses standard digital GPS from parallax. To overcome the problems of position accuracy with this model, we modified the design with commercial GPS and voice recognition unit. Figure 11(c) shows the modified unmanned robot, VizzBot-I.

![Figure 11(a). 3D Model of a Mobile Robot](image)

![Figure 11(b). Prototype Mobile Robot](image)

![Figure 11(c). Modified Version of Unmanned Mobile Robot VizzBot-I](image)
6. ALGORITHM SIMULATION FOR PATH PLANNING AND MODELLING OF OBSTACLES

This section is to show the simulation of our Algorithm for path planning and image processing which was implemented on a low cost, on board vision module. Because it’s not possible to visualize how the image is changing in the embedded controller. So the same algorithm is implemented in MATLAB to demonstrate the simulation.

As discussed in the previous section YUV color space is implemented on vision module for reduction in computation cost and to avoid the drastic effects of lightning conditions in the environment. The real time RGB image is captured from on board vision module and is shown in Figure 12(a).

Figure 12(a). RGB Image Showing the Working Domain

Figure 12(c). Modeling of 0’s and 1’s in the Working Domain

Figure 12(d). Edges of Obstacle and Dark Regions

Figure 12(a) shows a colored image acquired from onboard vision module showing the workspace where the mobile robot has to find a path. Note that in this figure the obstacles and the shades of these obstacles are pointed out. Figure 12(b) shows the working domain in YUV color space. In this figure, the shades of the obstacles are converted to dark region with the aim of converting them into obstacles. Using edge detection technique, edges of the obstacles and also that of the dark regions are determined and shown in Figure 12(d).

As discussed in [1] the potential function used in this design is

\[ U(q) = kp \left[ I(q) - I^2(q) / 2 \right] \]  

(1)

where \( q \) is a spatial point in the robot workspace which defines the robot's current position, \( I(q) \) represents the intensity of light at the point \( q \) and \( U(q) \) is a positive and continuous function which becomes zero when \( q = q_{goal} \). The intensity values here represent the values pertaining to 0's and 1's. Here \( kp \) is the position gain, if using a conventional servo. The force generated by this potential field is given as the artificial force thus acting on the robot is then given as

\[ F = kp \left[ 1 - I(q) \right] \]  

(2)

Thus the robot moves under this artificial force and the force becomes one when the intensity at a spatial point is less than the threshold value and at a lighted region, the force becomes zero. The planning of the path of the robot is then done such that the robot follows the zero-force path.
Plotting of potential function: The points generated in the edge detection are projected on the potential function, the results of which are shown in Figure 12(d). In this figure, the plain region represents a free space where the robot is allowed to navigate. The robot treats the plain region as the attractive potential field and the peaks as the repulsive potential field. With respect to the real image acquired, the peaks represent the regions having intensity values less than the threshold value. The plain region shows the areas having intensity values greater than the threshold value. Figure 12(e) shows the repulsive potential field which represents the obstacles and the dark regions. Modeling of the dark regions is discussed in section 5. Figure 12(f) shows the obstacles and the free space in the working domain of the image shown in Figure 12(a). The red circles, in this Figure 12(f) are the obstacles and the white background is the free space where the robot can navigate.

7. ALGORITHM IMPLEMENTED IN VIZZBOT-I BRAIN

Figure 13 explains the algorithm implemented in the Brain of VizzBot (P81V51RD2). This takes the obstacle information from camera and direction information from commercial GPS (voice recognition unit) in real time at every instant and computes the direction of travel using path planning algorithm discussed in [1] so as to reach the destination in the optimal path.

8. CONCLUSION

In this paper, a novel approach of interfacing commercial GPS with unmanned mobile robot is discussed by interfacing HM2007 with the TomTom GPS. The major problems concerning the
position accuracies of using GPS in unmanned mobile robots can be resolved by using the suggested technique. This technique resulted in position accuracy of better than 0.1 m, which means the position accuracy of GPS based unmanned mobile robot was improved by 50 times. Furthermore, the path planning algorithm in [1] was modified and improvised for VizzBot-I to have optimum performance. This was particularly suitable for cameras having no infrared sensors in which case the robot may not distinguish between bright and dark regions. In the absence of this distinction, the robot cannot view the obstacles in the dark regions. Additionally, in cases where the robot moves into the dark regions then it may not clearly visualize any of the obstacles in the environment and hence the robot may collide with the obstacles. Thus, a potential function is used to model varying intensity of light in the workspace such that the robot moves only along the lighted areas of the workspace and, at the same time models the obstacles and tread an optimal path. This approach significantly improved the navigation by allowing the robot to avoid deadlock conditions. A GPS interfaced with voice recognition unit is configured with the VizzBot-I platform to have the direction information for guiding the robot to the final destination, which was successfully demonstrated.

9. REFERENCES


[26] B.Faverjon and P.Tournassoud. A local based approach for path planning of manipulators with a high number of degrees
