Fusing Radar and Vision for Detecting, Classifying and Avoiding Roadway Obstacles

Dirk Langer and Todd Jochem
The Robotics Institute,
Carnegie Mellon University, Pittsburgh, PA 15213
Phone +01-412-268-2000; Fax +01-412-268-5570; E-mail langer+@ri.cmu.edu, tjochem@ri.cmu.edu

Abstract

This paper describes an integrated MMW radar and vision sensor system for autonomous on-road navigation. The radar sensor has a range of approximately 200 metres and uses a linear array of receivers and wavefront reconstruction techniques to compute range and bearing of objects within the field of view. It is integrated with a vision based lane keeping system to accurately detect and classify obstacles with respect to the danger they pose to the vehicle and to execute required avoidance maneuvres.

1 Introduction

Most on-road research has focussed on a single part of the navigation problem, such as lane keeping, obstacle detection, or route planning. This paper presents work on combining multiple modalities into an integrated obstacle detection, classification, and avoidance system.

In the context of an autonomous automobile driving at moderate or high speeds, an obstacle detection sensor with a fairly long range is needed (refer also to [6]). This sensor needs to be able to operate robustly under adverse weather conditions, when long range visibility is poor and provide object location at a sufficient longitudinal and lateral resolution and data rate. A MMW radar sensor was designed to accomplish this task.

For a vehicle driving on a road, objects in the vehicle's driving lane and neighboring lanes are of immediate concern. It is generally not possible to tell whether a detected target would interfere with the vehicle's planned motion in cluttered environments and/or on curved roads using only radar data. Additional road geometry information is required. Here, lane position information from RALPH [4] is used to classify the degree of danger that an obstacle poses to the vehicle and to plan appropriate avoidance maneuvres when required. RALPH's output is then used to move the vehicle along a safe trajectory around the obstacle, possibly executing a complete lane change.
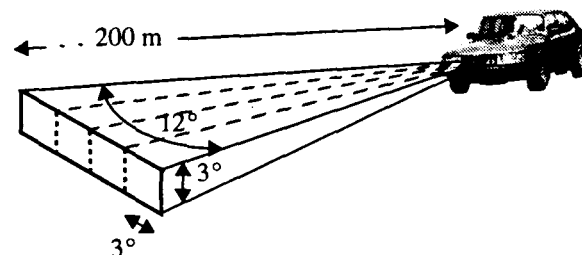
2 Radar Sensor Design Concept

Figure 1 Sensor Geometry

The geometry of the radar sensor is shown in Figure 1. We choose a vertical field of view (VFoV) of 3° as it provides a good compromise between good obstacle coverage in the vertical direction and avoiding false measurements due to ground reflections and returns from road signs or other structures located overhead. At longer ranges the ground (road) will reflect specularly. The current design shows a horizontal field of view (HFoV) of 12°, which is divided into four or more angular resolution cells (See also section 2.2). Assuming an average highway lane width of 4 metres, the sensor covers one lane at a range of 19 metres and three lanes at a range of 57 metres. At a range of 95 metres the sensor covers an area of 20 m by 5 m.

We decided on the approximate required maximum range of the sensor by the following reasoning:

In order to stop for a stationary object at highway speeds of 100 km/h (~ 65 mph) a detection distance of at least 55 metres would be needed (assuming $a = -7$ m/s^2 ≈ -0.7 g and neglecting reaction time). Since in autonomous mode, the vehicle needs to react to objects moving at considerably different

relative speeds, the maximum range of the sensor for detecting vehicles is designed to be about 200 metres. Humans have a much lower radar cross section (0.2 - 2 m^2) as compared to vehicles (10 m^2) and can therefore be detected only at a much smaller maximum range. For the sensor system, this would be about 60 metres. However, this does not pose a problem as humans are assumed only to be present in a road environment where vehicles also move considerably slower, i.e. 50 km/h (~30 mph).

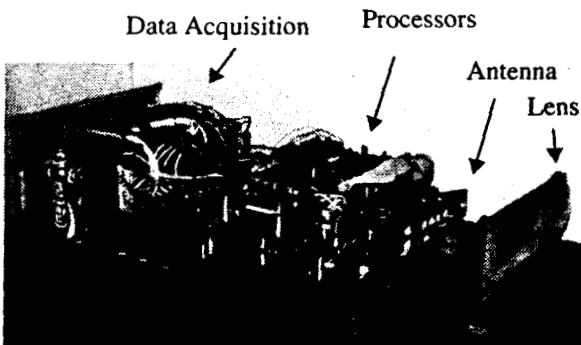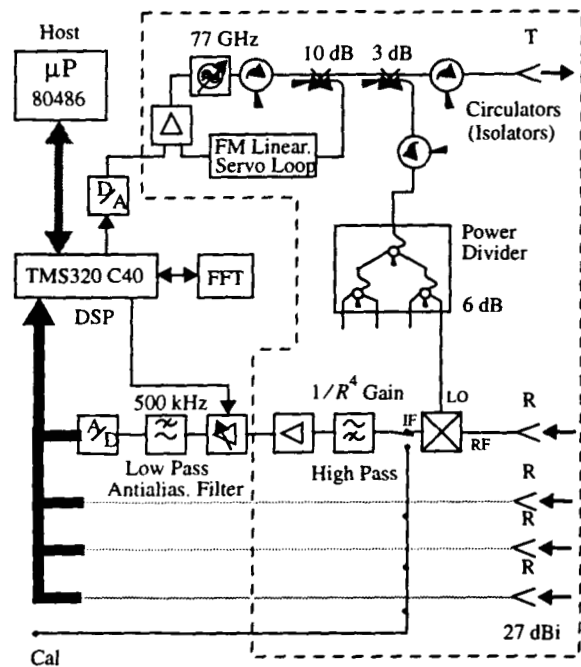## 2.1 Radar Specifications

Figure 2 FMCW Radar Block Diagram and Picture

For simplicity of design and better sensitivity, we decided to operate the radar as a Frequency Modulated Continuous Wave (FMCW) system. A block diagram of the design is shown in Figure 2.

The carrier frequency is at 77 GHz. The corresponding wavelength is in the mm-Range which results in better resolution of scene details than at lower frequencies. Also, above 30 GHz, reflection from vegetation and foliage is more significant. Attenuation at these frequencies is larger than at lower ones. However, this also keeps the maximum range of the radar relatively short and thus reduces interference.

A carrier frequency in the millimeter wave range also allows larger frequency sweeps for better range resolution and limits interference with existing commercial radio frequency transmitters. Because the transmitted signal is spread over a large bandwidth, an FMCW radar is quite robust against interference from other sensors of the same type. Interference can be further reduced by using a coded FM wave form. For a simple radar system the maximum range resolution is related to the bandwidth by the following relation:

$$\delta R = \frac{c}{2B} \qquad B = 2\Delta f \qquad (1)$$

The frequency sweep $2\Delta f$ of the radar is 300 MHz, which therefore results in a range resolution of 0.5 m. For a maximum range of $R = 200$ m and a desired maximum intermediate frequency (IF) of $f_{IF} = 500$ kHz, the modulation frequency f_m is given by equation (2) and is calculated to be 625 Hz for a triangular modulation and
1.25 kHz for a sawtooth modulation:

$$f_m = \frac{c \cdot f_{IF}}{8\Delta f\, R} \qquad \text{(Bipolar Triangle)} \qquad (2)$$

If the FM waveform used is triangular, maximum unambiguous range is 1200 km and can be calculated from equation (3):

$$R_m = \frac{c}{2f_m} \qquad (3)$$

In order to be able to detect vehicles and people within the given range, transmitter power of the radar system is about 30 mW, as can be calculated from the *radar equation* (see [3]).

Range and angular information of targets are obtained by Fast Fourier Transform (FFT) or similar spectral analysis methods as described in the following section.

2.2 Data Processing

The output of the mixers on the receiver channels of the radar sensor is a mixture of different frequencies (see Figure 2 and Figure 3). Each discrete frequency corresponds to range to a target. Therefore the range to a target can be obtained efficiently by using a FFT.
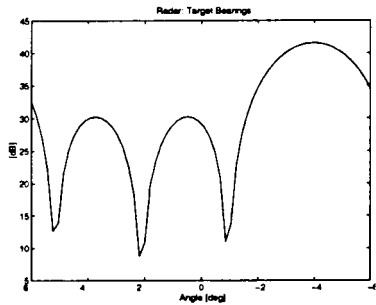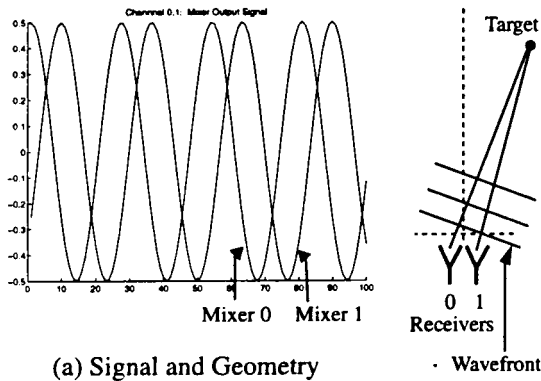

(a) Signal and Geometry

(b) Bearing (Signal Phase)

Figure 3   Radar Simulation of target at (30 m, −4°)

Angular bearing of the target is obtained by digital wave front reconstruction and beam forming. This involves two consecutive FFTs along the time and space dimension of the signal. The space dimension is in this case the four receiver elements that are used to sample the incoming wave front at discrete points in space. The basic idea is shown in Figure 3 for a simulated target at (30m, −4°); only two out of the four receivers are shown. A parallel wavefront reflected from a target, offset from the major sensor axis, is incident at an angle on the receiver array. As can be seen in this case, the wavefront arrives slightly earlier in time on receiver 1 than on receiver 0. Thus there is a small difference in range which results in a slightly different frequency output between receivers. This small frequency difference cannot be resolved by the FFT in range, but it shows up as a phase difference

in the time signal (Figure 3 (a)).

The first FFT along the time dimension of the receiver output signal now gives target range as mentioned above and relative phase information. The second FFT along the space dimension performs a cross correlation that is a measure of the phase (time) lag between the signals in each receiver channel. The maximum peak in the correlation indicates the phase (time) lag and thus the bearing of the incoming signal (Figure 3(b)).
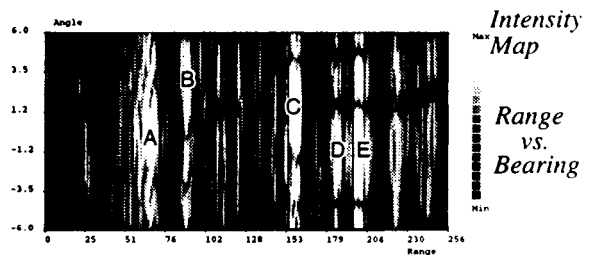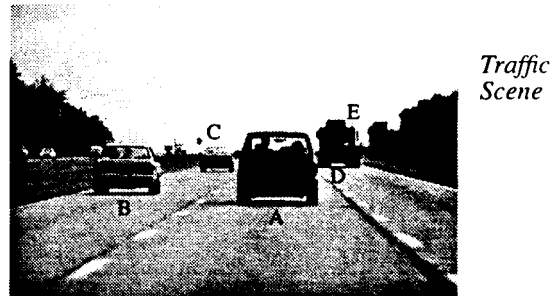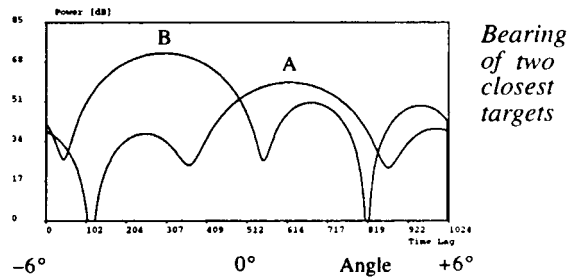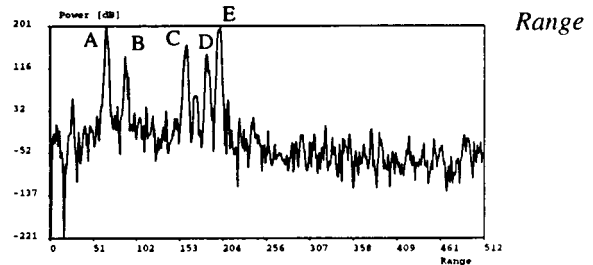

Range

Bearing of two closest targets

Traffic Scene

Intensity Map

Range vs. Bearing

Figure 4   Plot of data obtained on Highway

A more detailed description of the method is given

in [1] and [2]. Since the radar has four receivers, we obtain four angular resolution cells as shown in Figure 1. In order to determine target bearing unambigously within the given sensor's field of view, the following condition for the receiver antenna spacings Δx_i needs to be obeyed:

$$\Delta x_i \; < \; \frac{1}{\sin(2\alpha_E)} \cdot \lambda \qquad (4)$$

where $2\alpha_E = 12°$ is the unambigous range, i.e. the horizontal field of view.

The Fourier Transform is limited in resolution by the number of data points given. This means for the bearing dimension that one resolution cell is $12°/4 = 3°$, since we have four receivers in the spatial dimension. It should be noted that therefore if two targets are at the same range, but their bearing is less than $3°$ apart, they cannot be resolved anymore and are merged into a single target lobe.

It should be noted that here resolution is the ability of the system to distinguish between two separate targets that are close together, whereas accuracy is the absolute accuracy with which a single target position can be determined. An increased resolution is generally desirable as it leads to an improved Signal-to-Clutter S/C ratio. This improvement happens because ground and rain clutter generally fill an entire radar resolution cell, but the dimensions of a typical radar target are usually much smaller.

Figure 4 shows plots of range and bearing and the corresponding scene picture of real data obtained with the sensor on a highway. Note that the bearing data appears quite smooth because only 4 data points are available.

3 Object Classification

Detected targets are now kept in an object list (or local map) which is updated by successive data frames from the radar sensor. Each object has the following set of parameters or annotations associated with it:

{Position (x,y), Amplitude, History, Decay, Lane}

Position indicates the x-y position of the object with respect to the vehicle centre. Amplitude is the target strength. History counts the number of times this particular object has been detected and tracked in successive frames. Decay indicates when an object will be deleted from the list if it has not been seen for a successive number of times. Lane indicates the road lane or off road location of an object with respect to the vehicle's driving lane.

Objects are tracked through successive frames within a specific window area and by using information about vehicle position changes. The size of the tracking window depends on the expected relative maximum speeds between objects and the frame update rate of the sensor.

In order to assess the danger an object poses to the vehicle's motion, we need to know the location of an object with respect to the vehicle's driving lane. This is a simple problem on a straight stretch of road, using only radar sensor data. However, on curved roads in cluttered environments it cannot be solved without additional information about road geometry (Figure 5).
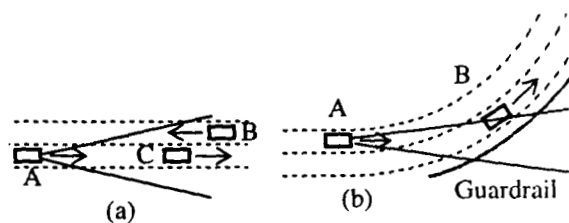

Figure 5 Traffic situations leading to potential false alarms

This information is provided by the RALPH lane tracking system described in section 4. For classification purposes the road lane centre of the vehicle's driving lane is given as a set of points $P_i(x_i, y_i)$ and the road curvature $1/r_i$ between P_i and P_{i-1}, where P_0 is at the vehicle centre. Hence the road lane can be represented as a set of straight line and circular arc segments as shown in equation (5) for segment i.

$$y = m_i \cdot x + c_i \quad \text{or} \quad (x - h_i)^2 + (y - k_i)^2 = r_i^2 \qquad (5)$$

A straight line or circular arc is chosen depending on the road curvature of that particular segment.

Given equation (5) and assuming an average highway lane width of 4 metres or the lane width estimate given by RALPH, we can now determine whether an object is inside or outside the current lane boundary and its distance to the boundary. The Lane parameter for each object in the list is then updated accordingly.

The information in the local obstacle map can then be used by the Navigator in order to take appropriate actions, depending on distance to closest obstacle in the vehicle's driving lane and obstacles in adjacent lanes (see Figure 7 and section 4).

The above road representation also allows the

integration of other sources for road geometry information with further look ahead distances such as a digital road map.

4 RALPH

RALPH is a vision based adaptive system which can learn the current road features. It uses a weak model of road geometry and image reprojection to extract and adapt to the relevant features for driving.
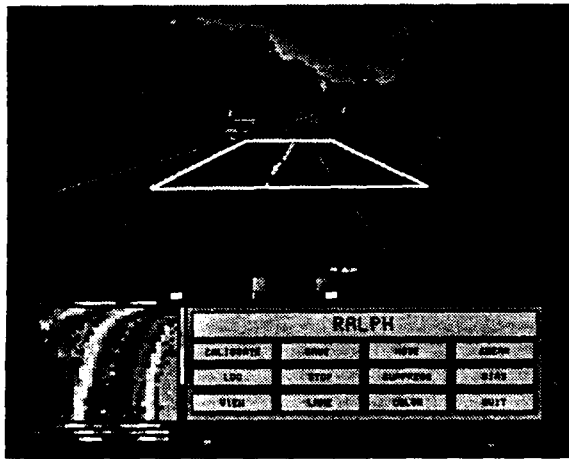

Figure 6 RALPH display for typical highway scene

In order to locate the road ahead, RALPH first resamples a trapezoid shaped area in the video image, much like a birds-eye virtual camera, to eliminate the effect of perspective. RALPH then uses a template-based matching technique to find parallel image features in this perspective free image. These features can be as distinct as lane markings, or as subtle as the diffuse oil spots down the center of the lane left by previous vehicles. RALPH rapidly adapts to varying road appearance and changing environmental conditions by altering the features it utilizes to find the road. This rapid adaptation is accomplished in under one second, without human intervention [5].

Since RALPH can exploit any visible features running parallel to the lane, instead of relying exclusively on the presence of distinct lane markings, it can operate in a wider variety of situations than previous road following systems. In one experiment, called "No Hands Across America", RALPH drove the Carnegie Mellon Navlab 5 test bed vehicle 98% of the 2850 mile journey from Washington, DC to San Diego, CA [5]. During the trip, RALPH drove at an average speed of 63 m.p.h., in conditions which included bright sunlight, dusk, rain and nighttime. During one stretch in Kansas, RALPH drove continuously for 69 miles without the safety driver touching the steering wheel.

5 Lane Changing

Under normal driving conditions, the Control Flow Coordinator (CFC) takes RALPH's computed vehicle lateral displacement from the center of the lane and its estimated road curvature value and passes them to the vehicle controller which attempts to recenter the vehicle in the driving lane.
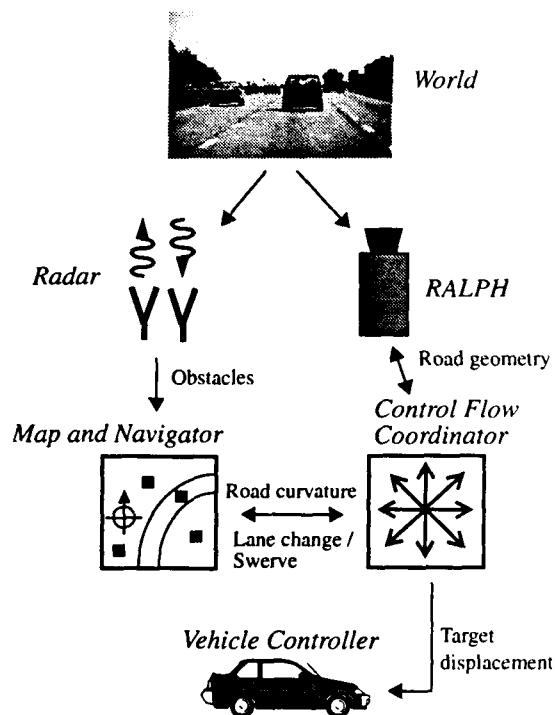

Figure 7 Data and signal flow diagram

When the CFC receives a signal from the Navigator that a lane change maneuvre is necessary, it first determines which direction of lane change should be executed. To do this, it uses RALPH's estimate of the current lane layout - whether the vehicle is in the left, middle or right driving lane. If the vehicle is in the left lane, the lane change direction will be to the right. The opposite is true when the vehicle is in the right lane. In the cases where the vehicle is in a middle lane, or when RALPH does not reliably know which lane the vehicle is in, a lane change direction opposite of the last one executed is used.

After the lane change direction has been determined, the CFC incrementally adjusts the target

vehicle offset away from the center of the driving lane (offset of zero) and toward the destination lane. When the controller attempts to align the vehicle at this new offset the effect is to move the vehicle towards the destination lane. The controller servos to the new offset using RALPH's estimate of the current lateral displacement from the center of the lane as ground truth. In the case of a right-to-left lane change, the target offset becomes incrementally more negative.

The increment that is added to the target offset at each step of the lane change is computed so as to maintain a constant lateral velocity. The target lateral velocity that is used is typically between 0.4 and 0.6 metres per second. These bounds were computed from previous research on a similar system which showed that they yielded comfortable, controlled lane transitions with a duration between 6 and 8 seconds [4].

This incremental adjustment procedure is used until the vehicle reaches the midpoint between the source and destination lanes. At this point, because the features that RALPH uses to track the source lane begin to leave the image, the CFC instructs RALPH to stop looking at those old features, and begin looking for new features in the destination lane. The CFC informs RALPH of the approximate lateral location of the center of the destination lane which RALPH uses to lock onto new features appropriately.

After RALPH has locked onto the destination lane features, the output displacement which it sends to the CFC changes sign. This occurs because the reference point for the displacement has changed from the center of the source lane, which for a right-to-left transition was to the right of the vehicle, to the center of the left lane.

When the CFC senses this sign change, it instructs the vehicle controller to reset its internal control loop parameters and begin processing incoming values with respect to the new reference frame. For the right-to-left lane change, the CFC continues incrementally decreasing the target offset, based on the constant lateral velocity model, until the vehicle has reached the center of the destination lane. At this point, normal lane keeping resumes.

6 Conclusions

The described integrated radar and vision sensor system demonstrates the ability of robust high speed autonomous vehicle navigation and obstacle detection. This is achieved by fusing data from two different sensor systems, radar and vision, resulting in an improved vehicle navigation performance in a variety of different traffic and road scenarios. First results were obtained from the real system in a highway environment. Apart from being used in an autonomous navigation system, the sensor system would also be capable of providing relevant information about the local traffic situation to an intelligent cruise control (ICC) or a human driver. As an improvement over previous systems, it is able to operate on a highway as well as more cluttered environments such as rural roads, where currently the largest number of fatalities occur.

Acknowledgments

References

[1] M. Bockmair. *Mikrowellenabbildung im Nahbereich*. Ph.D. Dissertation, TU Muenchen, 1989.

[2] P. Cao, J. Cuschieri and R. Sudhakar. A High Resolution 3-D Sonar Imaging System For Autonomous Underwater Vehicles. In *Proc. of Oceans '91*. Honolulu, HI. October 1991.

[3] Jerry Eaves and Edward Reedy. *Principles of Modern Radar*. Van Nostrand Reinhold Company, 1987.

[4] Jochem, Todd M., Pomerleau, Dean A., and Thorpe, Charles E. Vision Guided Lane Transition, In *Proc. IEEE Symposium on Intelligent Vehicles*, September 1995, Detroit, Michigan, USA.

[5] Pomerleau, Dean A. and Jochem, Todd M. "A Rapidly Adapting Machine Vision for Automated Vehicle Steering," IEEE Expert: Special Issue on Intelligent System and their Applications, pp. 19-27, April 1996.

[6] M. Rozmann, M. Lange, J. Detlefsen. Collision Avoidance Radar Using Wavefront Reconstruction. In *Proc. 3rd PROCHIP Workshop*, Paris, May, 14th 1990, pp 251-256.