Perception for collision avoidance and autonomous driving

Romuald Aufrère, Jay Gowdy, Christoph Mertz[*], Chuck Thorpe, Chieh-Chih Wang,
Teruko Yata
The Robotics Institute, Carnegie Mellon University,
5000 Forbes Avenue, Pittsburgh, PA 15213, USA

1   Abstract

The Navlab group at Carnegie Mellon University has a long history of development of automated vehicles and intelligent systems for driver assistance. The earlier work of the group concentrated on road following, cross-country driving, and obstacle detection. The new focus is on short-range sensing, to look all around the vehicle for safe driving. The current system uses video sensing, laser rangefinders, a novel light-stripe rangefinder, software to process each sensor individually, a map-based fusion system, and a probability based predictive model. The complete system has been demonstrated on the Navlab 11 vehicle for monitoring the environment of a vehicle driving through a cluttered urban environment, detecting and tracking fixed objects, moving objects, pedestrians, curbs, and roads.

2   Robotics Technology and Intelligent Vehicles

During the decade of the 1990's, researchers from automotive engineering, from robotics, and from allied fields came together to start building intelligent vehicles [1,2,3,4]. These vehicles use sensors, actuators, processors, and communications systems to either drive automatically, or to monitor a human driver and assist in navigation or warn in case of a developing dangerous situation.

The field was driven by several market demand factors. First is an increasing interest in safety. While the accident rate has decreased steadily for 40 years, the total miles traveled have increased, with the net result that the number of road fatalities in the US has stayed approximately constant, averaging 40,000 / year [5]. Furthermore, there is an increasing sense that vehicles have become safer (air bags, anti-lock brakes, crumple zones); roadways have become safer (the Interstate highway system); societal pressures have become stronger (mandatory child safety seats, serious enforcement of drunk driving laws) but the drivers themselves have not necessarily improved. Currently, in the US, 70% of all crashes are primarily caused by the drivers, and in another 20% driver error is a major contributing factor [5,6].

Beyond safety, the next major driving factors have to do with congestion and pollution. In the US, traffic continues to grow at 4% / year, while with the completion of the Interstate highway system, major new road construction greatly lags traffic growth. Computer-controlled vehicles promise to smooth out traffic flow, permitting more vehicles to travel in

---

[*] Corresponding author,
E-mail address: cmertz@andrew.cmu.edu

a given lane. With conventional vehicles, the best traffic flows are in the vicinity of 2,000 vehicles / lane / hour; at 100 kph this works out to 50 m / vehicle. Of course, real traffic does not flow with even separations between vehicles; instead, traffic bunches and gaps. Smoothing out the traffic flow, with an even spacing of 25 meters, would of course double the throughput, without having to pave more lanes.

Finally, the push for intelligent vehicles is driven by the pressures of an aging society. Older people may, for instance, lose flexibility in their neck, making it harder to check before changing lanes; intelligent sensors that monitor the blind spot may allow seniors to drive safely for more years.

In addition, the military has increasing interest in intelligent vehicles. Scout vehicles typically operate in front of the main force, and are the first target of the opposition. The US Army, in particular, has long desired to have automated scouts that can investigate hazardous areas without putting soldiers in harm's way.

At the same time as the demands for intelligent vehicles were increasing, the technologies to build them were becoming available. Advances have been driven by the military, aerospace, and electronics; even the entertainment industry has helped, such as the proliferation of camcorders driving down the cost of CCD video sensors.

From an organizational standpoint, the field rapidly became organized with all the usual trappings:
- In 1990 the first Intelligent Vehicles workshop was held, sponsored in part by IEEE. This has now become an ongoing series of workshops [7,8,9,10,11].
- In March 1991 the trade organization ITS America was chartered by the government of the US as an official Utilized Federal Advisory Committee; it now has more than 600 member companies and organizations.

The biggest public event of the field was the Automated Highway Demonstration, held in 1997 [4]. The National Automated Highway Systems Consortium (NAHSC) was composed of representatives from the automotive industry (General Motors, with associates Honda and Toyota); construction (Bechtel and Parsons Brinkerhoff), state and federal agencies (CALTRANS and US DOT Federal Highways Administration), electronics (Delco), systems engineering (Hughes, Lockheed Martin), and academia (Carnegie Mellon, Berkeley). The AHS Demonstration showed 22 vehicles, carrying passengers, driving autonomously over a section of I-15 north of San Diego[12,13,14,15,16].

Parts of this technology are now coming on the market. Adaptive Cruise Control is now available on a few high-end passenger cars; this uses a radar or laser sensor to detect other moving vehicles in the lane ahead, and automatically adjusts the cruise control set speed to maintain a safe distance. ACC is sold as a convenience feature, rather than a safety feature, because it is currently not capable of distinguishing stopped vehicles in its own lane on a

curving road. Other sensors available on vehicles include side-looking sensors on large trucks to cover their blind spot; and prototype sensors based on forward-looking vision to detect drowsy drivers who are starting to drift out of their lane.

A common feature of all of the above systems is that they are designed for the open highway. This is an appropriate first step, but only addresses part of the problem. A controlled-access roadway is designed for easy perception and control, and should only have motor vehicles on it under normal circumstances.

In contrast, urban environments (where a large portion of driving actually takes place) are characterized by clutter, both fixed objects (lamp posts and stop signs and fire hydrants and newspaper vending machines and …) and by moving objects (cars and bicycles and pedestrians and dogs and …). As the problems of highway driving begin to be solved, there is a new push just starting in the research community on the particular needs of urban driving [17,18,19,20].

3   The need for 360 degree safeguarding

Robot driving has concentrated on forward-looking sensing, for road following and obstacle detection [21,22,23,24,25,26,27]. This is an appropriate first step, but real deployment of mobile robots will require additional sensing and reasoning to surround the robot with safeguard sensors and systems. Our group is currently building short-range sensing to surround vehicles to improve the safety of robotic and human-controlled vehicles.

In the civilian context, our focus is driver assistance for transit buses [28,29,30]. Buses drive at relatively slow speeds, in crowded urban environments. One of the most frequent types of accidents in transit buses is side collision, where the driver does not have adequate awareness of objects near the bus, then turns too sharply and sideswipes a fixed object or (less often) hits a pedestrian. Preventing these accidents requires short-range sensing along the side and front of the bus, detecting fixed objects, detecting and tracking moving objects, predicting the motion of the bus itself, and alerting the driver through a suitable driver interface. Moreover, the assistance system has to discriminate between safe and unsafe situations, which is especially difficult in the cluttered environments the bus is driving in. For example, pedestrians standing on the curb can be considered safe, but pedestrians in the road cannot.

In the military context, our focus is short-range sensing for full automation of scout vehicles [31]. An autonomous vehicle moving through a cluttered environment, such as a forest, may need to move between objects (e.g. trees) with very little clearance on either side of the vehicle. The conventional approach is to sense trees with forward-looking sensors, enter those trees into a map, and estimate the clearance as the vehicle moves forward and the trees move out of the field of view of the forward-looking sensor. If sensor data is noisy, or if the vehicle slips and slides in mud, the estimated clearance may be

incorrect. It is better to directly and continuously sense nearby objects all along the side of the vehicle as it moves through the forest.

Both civilian and military vehicles, and both driver assistance and full automation, need to pay special attention to moving objects, and particularly to humans. People move in unpredictable ways. Seeing a person with a forward-looking sensor, or having the driver note the position of a person in front of a bus, is no guarantee that the person will remain safely out of the vehicle's way. Thus, we pay special attention to detecting moving objects, classifying them if possible as people, continuously tracking their motion, and attempting to predict their future actions.

4 Testbed Vehicle

Figure 1: Navlab 11 testbed showing cameras, GPS, Laser Scanners, and a Laser Line Striper (one arrow points to the laser, the other to the camera). On the left side of the vehicle is another camera and Laser Scanner, positioned symmetrically to the ones on the right side.

Our testbed vehicle for this work is the Navlab 11 (Figure 1). It is equipped with:
- motion sensors (IMU, GPS, differential odometry, compass, inclinometer, and angular gyro);
- video sensors (6 video cameras, 3 in front, one on top, and two looking along the sides);
- ladars (3 Sick single-axis scanning laser rangefinders mounted in various positions, typically one looking forward and two along the sides);

- a light-stripe rangefinder typically mounted looking for the curb on the right;
- 5 500-MHz Pentium computers;
- high-accuracy time synchronization;
- various other sensors and actuators.

5   Perception Modules

We have built, tested, and integrated a number of perception modules, designed specifically for short-range sensing.

Optical Flow: Our work with optical flow uses an Omnicam and optical flow tracking to detect nearby objects from a moving vehicle [32,33]. The Omnicam is a video sensor with a hemispherical mirror, so a single image captures a 360-degree view of the world (Figure 2).
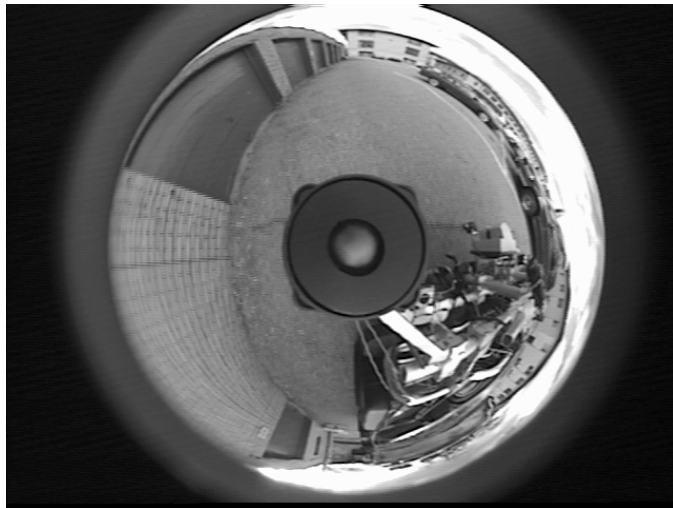

Figure 2 Raw image of the Omnicam.

In software, we unwarp the image, detect vertical edges, and then track those edges from frame to frame as the vehicle moves (Figure 3).
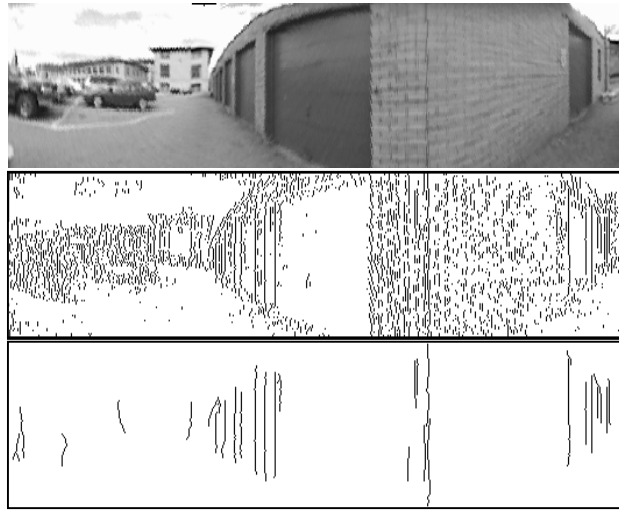
Figure 3 Processed Omnicam image: a) unwarpped image, b) vertical edges detected in the image, c) tracked edges.

We use a Kalman filter to track each of the detected objects, to update the position estimate of the objects with time. The result is a map of all the detected objects, suitable for obstacle avoidance and map building (Figure 4). This subsystem has not yet been integrated into the full data fusion path.
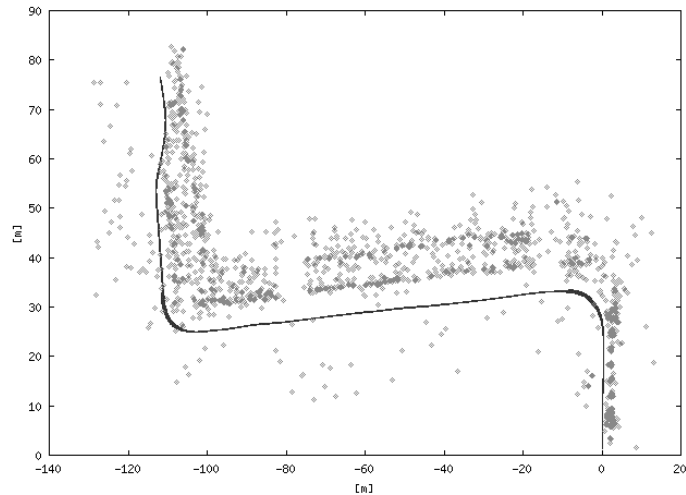
Figure 4: Map of tracking result. The solid black line is the path of the vehicle, the gray dots are tracked features.

Light-stripe scanner: We have designed, built, and integrated a light-stripe sensor as a proof-of-principle of a no-moving-parts range sensor, designed for use outdoors at modest

ranges [34]. We tested the sensor on board a transit bus operating in the city of Pittsburgh under a full range of weather conditions. The idea of light-stripe sensing is well-known (Figure 5):
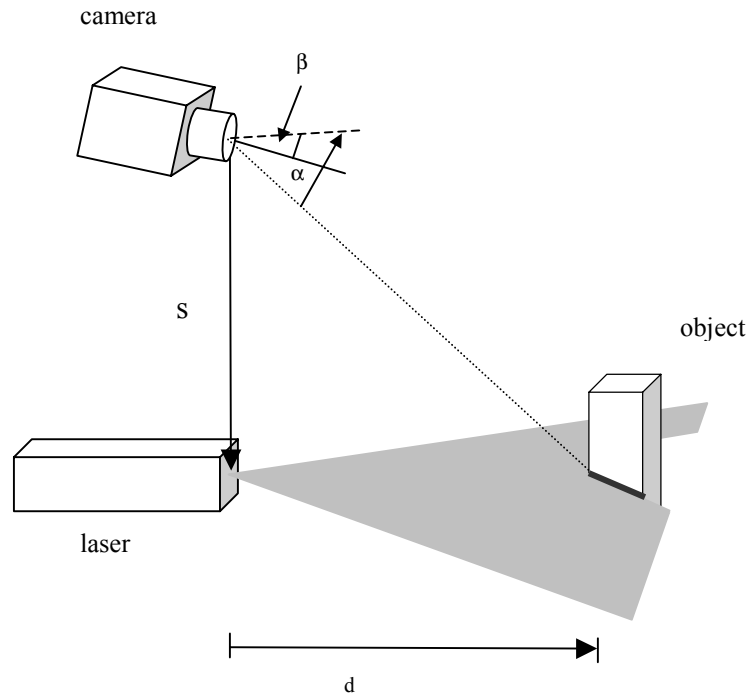
**Figure 5: Geometric configuration of the laser and camera of the laser line striper.**

A laser shines a plane of light on the scene while a camera located at a distance from the laser observes the scene (Figure 5). The camera sees the intersection of the laser light with objects at an angle α. The distance between the laser and the object can then be calculated by triangulation:

$$d = s \cdot \cot(\alpha)$$

The novel additions we have made to the process enable range sensing in full sunlight outdoors. First, we use a lens filter on the camera, so that it only sees light in the same wavelength as the laser illumination. Second, we synchronies the camera shutter with a pulsed laser. That allows much higher instantaneous power output, while still having a total integrated emitted energy low enough to keep the laser eye safe. For most outdoor conditions these two methods are sufficient to eliminate the background of daylight. But on a bright day, sunlight still causes erroneous measurements. They can be eliminated either through noise filtering algorithms or by taking a second background image without the laser on and subtracting it from the first.

Curb detection and tracking: In use, our side-looking light stripe range sensor measures the vertical profile of the road to the right of the vehicle. In our test vehicle the sensor is

mounted on the front (Figure 1), on a transit bus, it can be mounted inside the front bumper (Figure 6).
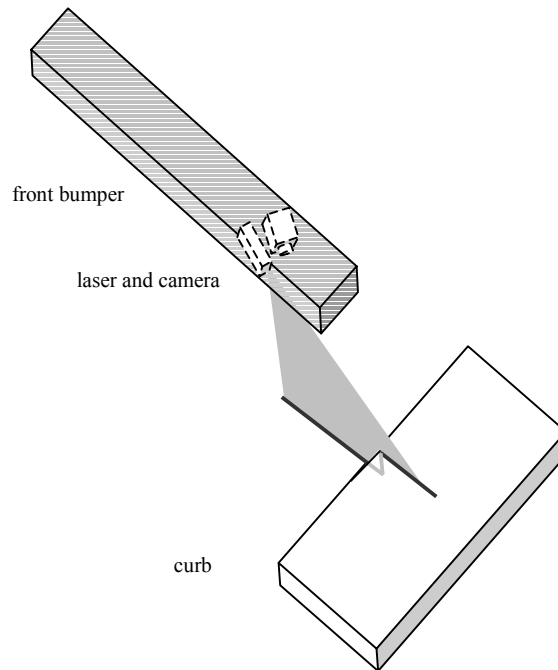
Figure 6 Schematic of the curb detector mounted inside the bumper of a transit bus.

When there is a curb present, the light stripe sensor can find the curb out to a range of approximately one lane width. We have tuned the software to detect the kinds of curbs commonly encountered in urban environments. We also built a Kalman filter to integrate curb measurements and vehicle motion models over time, to create a track of where the curb has been and an estimate of its current heading (Figure 7).

Curb detection in front of the vehicle: Once the curb has been detected and tracked, we project its location into the video image (Figure 7) acquired by the camera mounted on the rear right of the testbed looking along the vehicle (Figure 1). Given the location and direction of the curb in this image, an edge detector is used to search for the edge of the curb further in front of the vehicle. Figure 7 shows the result of the curb detection with the line striper, tracking of the curb alongside the vehicle, and detecting the curb in front of the vehicle using the image itself.

Figure 7 Image taken by the right forward-looking camera. Overlaid is the current curb position detected by the line striper (black dot), the curb position tracked 1 m alongside the vehicle, and the curb detected in front of the vehicle using this image.

The curb detection from the image gets complicated when an object is occluding the curb, *e.g.* if a person stands in front of it. Then the edge detector finds the edge of the pedestrian instead of the edge of the curb. This problem can be overcome if one takes the information of the laser scanner into account. The laser scanner can detect the objects in the scene. The areas in the image corresponding to the detected objects are excluded from the search of the edge of the curb. The result is, that the search will stop in front of the object (Figure 8).

Figure 8 Obstacle in front of the curb. The laser scanner detects the obstacle (white circles) and the search of the curb in front of the vehicle stops when it reaches the object.

Sick processing: We have been using a Sick single-axis ladar for map building, moving object detection, and ego-motion detection [35]. Since the Sick scans a single line across the scene, a horizontally-mounted Sick sensor maps a single plane of the scene. As long as the vehicle does not pitch or roll violently, and as long as the objects in the scene have adequate vertical extent, we can track those objects from frame to frame as the vehicle

moves. In Figure 9, the left-hand side shows the data from a single rangefinder mounted on the right of the Navlab11. The solid rectangle represents the vehicle itself; the points show range returns. The right-hand side shows the result of data matching. The long vertical line is the wall of a building that has been seen in multiple frames. The partial rectangle near the bottom is a parked car, that has been observed from the rear, side, and front, so a more complete representation has been built. The circles in the top of the frame show small objects that have been observed multiple times, but with changing locations; these are tentatively classified as pedestrians. The faint chain of circles trailing from the vehicle is the list of positions from which the vehicle made its observations.
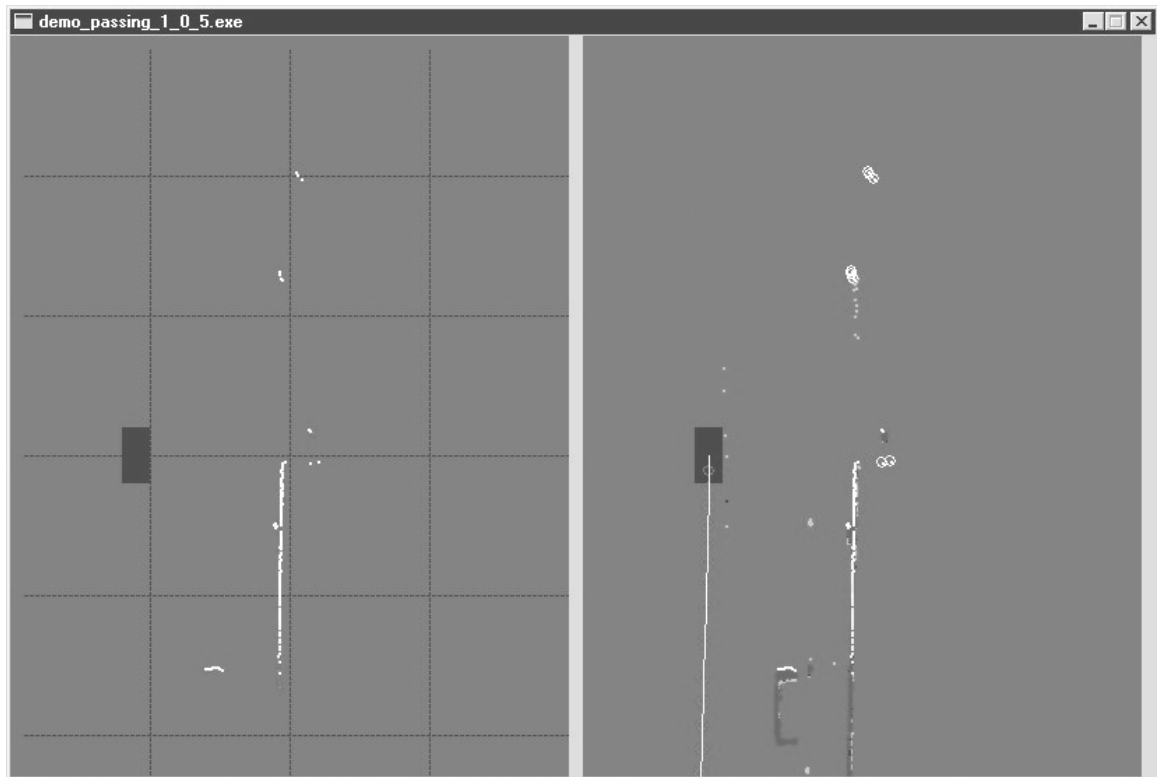
Figure 9: Sick scanner range data processing, showing a single scan (left) and multi-scan matching (right)

Fusion: The individual perception and interpretation modules all generate output that is fused into a consistent map representation. The data flow is shown in Figure 10. The "vehicle motion model", which keeps track of vehicle pose, is distributed to all modules. The other modules each process a variety of data, culminating in labeled objects flowing into the map.
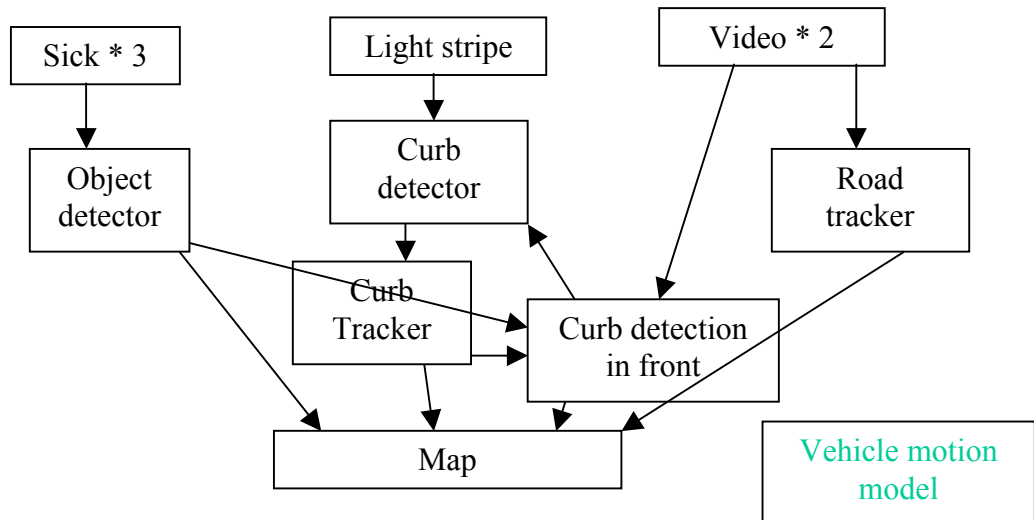
Figure 10: Data Fusion paths

6 Collision Prediction

Given the location of moving and fixed objects as collected in the map, and the heading of the vehicle, it is possible to calculate the time of collision for each object. However, since objects may change their motion, and the vehicle may not continue along its current arc, those calculations cannot be exact, they must take into account the probability distribution of likely future states of the testbed vehicle and other objects. This calculation is further complicated by different motion patterns for different kinds of objects: pedestrians on a sidewalk are less likely to step off the curb in front of a moving vehicle, and more likely to continue along the sidewalk. We have developed a probabilistic model of object motion that takes this kind of factor into account and generates a probability of collision as a function of time for each object in the environment. Rather than picking a fixed trajectory, we sample from the trajectories that the object could follow: faster or slower, turning more or less, and (for pedestrians) more or less likely to step off a curb and into the roadway. Each of these sampled trajectories is then examined to see if it causes a collision with the vehicle, and at what time. The fraction of trajectories that causes collisions is used as the probability of collision, as a cumulative function of time. High-probability imminent collisions are triggers for either an urgent driver warning (for driver assistance systems) or an immediate cue for the vehicle to take evasive action (for automated vehicles).

7 Demonstration System
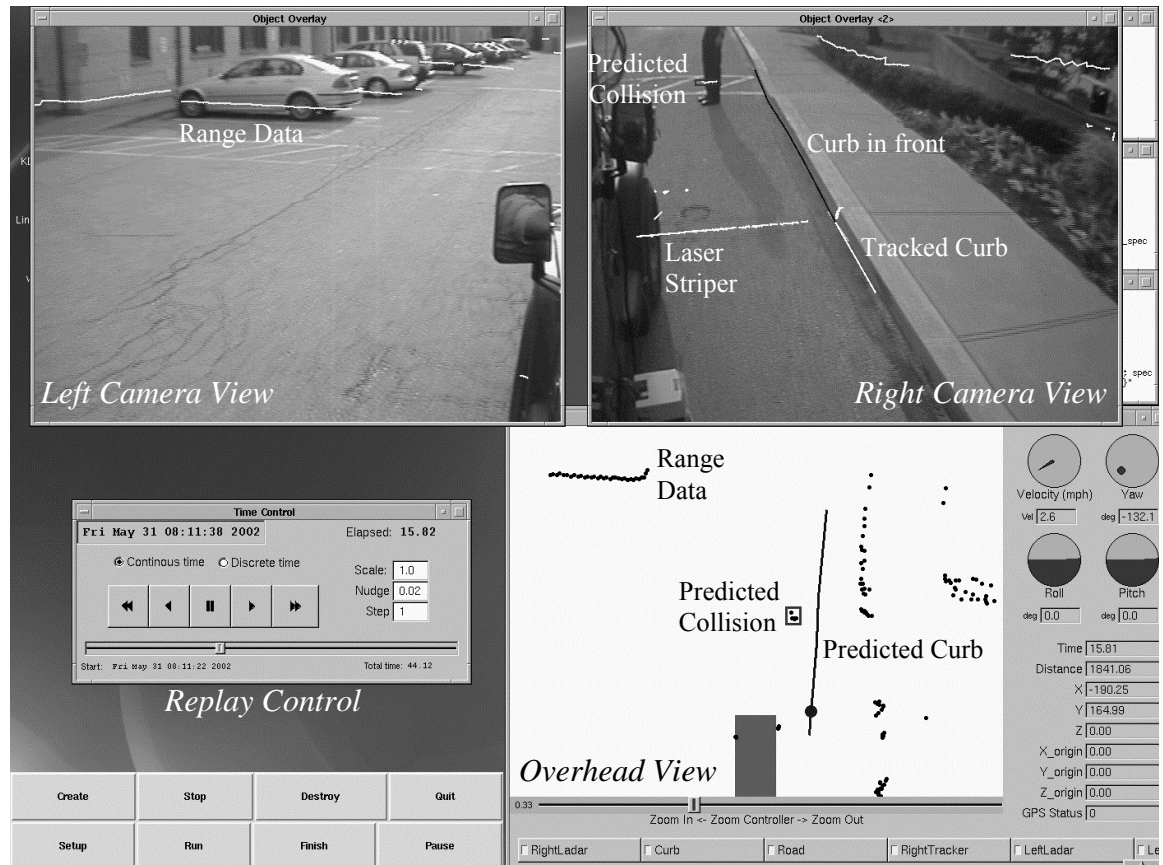
Replay of the integrated surround sensing system

Figure 11 Integrated demonstration system: left camera view with overlaid range data; right camera view with overlaid range data and curb location; and fused map.

All of these modules have been integrated into a demonstration system. The system does the following:

- Processes data in real time from two cameras looking along the sides of the vehicle, two laser scanners looking along the sides of the vehicle, the motion estimation sensors, and the side-looking curb tracker.
- Optionally, logs all the data with time tags (sub-millisecond accuracy) for later replay and analysis
- Finds objects in the environment
- Tracks the curb over time
- Overlays the detected objects and the curb on the video data
- Uses video processing to project the curb and road ahead of the vehicle, taking into account occlusions (as noted by 3-D objects projected onto the video image) to terminate tracking where the road edge is not visible

- Determines whether each object is in the roadway or safely on the side
- Estimates the potential of collision for each object

Figure 11 shows the complete system in operation. The displays on the bottom right show vehicle state. The left camera view, on the top left, shows the range data from the left Sick rangefinder and object detection system overlaid on the left camera. The top right shows the right camera view, with the right-hand Sick system data; plus the laser striper's current data, the detected curb, the track of the curb, and its future predicted position. The bottom right shows the combined map information.

8 Status and Ongoing Development

The systems as developed to date are good for demonstrations and proof of concept. The next steps are to mature the systems and deploy them on test vehicles for long-duration tests in real weather and real traffic conditions. Over the next year, the Navlab group will harden the electronics and improve the algorithms for full automation on robot vehicles, and for extended tests on transit buses.

9 Acknowledgements

References

[1] "Mobile Robots", C. Thorpe, International Journal of Pattern Recognition and Artificial Intelligence, Vol. 5, No. 3, August, 1991.
[2] "Mobile Robotics: Perspectives and Realities", C. Thorpe and M. Hebert, proceedings of International Workshop on Advanced Robotics and Intelligent Machines, April, 1996.
[3] "Intelligent Unmanned Ground Vehicles: Autonomous Navigation Research at Carnegie Mellon", M. Hebert, C. Thorpe, and A. Stentz Kluwer Academic Publishers, 1997.
[4] "Automated Highways and the Free Agent Demonstration" C. Thorpe, T. Jochem, and D. Pomerleau International Symposium on Robotics Research, October, 1997.
[5] National Transportation Statistics (www.bts.gov/btsprod/nts/).
[6] Fatality Analysis Reporting System (FARS, http://www-fars.nhtsa.dot.gov/) and the General Estimates System (GES, ftp://www.nhtsa.dot.gov/ges/).
[7] "Vision-based Vehicle Guidance", Masaki, Ichiro (ed) ,Springer-Verlag, 1991.
[8] Proceedings of IEEE Intelligent Vehicle Symposium, June 18-20, 2002, Versailles, France, Michel Parent.
[9] Proceedings of the 2001 IEEE Intelligent Vehicle Symposium, May 13-17, 2001, National Institute of Informatics, Tokyo, Japan.
[10] Proceedings of the 2000 IEEE Intelligent Vehicles Symposium, Dearborn, Michigan, October 3-5, 2000, Jim Rillings.
[11] Proceedings of the 1998 IEEE International Conference on Intelligent Vehicles, October 28-30, 1998, Stuttgart, Germany.
[12] "National Automated Highway System Demonstration: A Platoon System", Zhang, Wei-Bin, Proceedings of IEEE Conference on Intelligent Transportation Systems, Boston MA, November 1997.

[13] "Steps to Deployment: An Outline of Toyota's U.S. AHS Program", Andrews, Scott, in Proceedings of IEEE Conference on Intelligent Transportation Systems, Boston MA, November 1997.

[14] "Characteristics of the Honda-PATH AHS Team Vehicles for the 1997 AHS Feasibility Demonstration", Shladover, Steve and Empey, Dan, Proceedings of IEEE Conference on Intelligent Transportation Systems, Boston MA, November 1997.

[15]"The OSU Demo '97 Vehicle", Ozguner, Umit, et. al., Proceedings of IEEE Conference on Intelligent Transportation Systems, Boston MA, November 1997.

[16]"Infrastructure Diagnostic Vehicle", McKenzie, Patrick, in Proceedings of IEEE Conference on Intelligent Transportation Systems, Boston MA, November 1997.

[17] Proceedings of the IEEE 5th International Conference on Intelligent Transportation Systems, September 3-6, 2002, Singapore, Kan Chen and Fumio Harashima.

[18] Proceedings of the 4th International IEEE Conference on Intelligent Transportation Systems, Oakland, California, August 25-29, 2001, Pravin Varaiya and Hamed Benouar.

[19] Proceedings of the 3rd Annual Conference on Intelligent Transportation Systems, October 1-3, 2000, Dearborn, Michigan.

[20] Proceedings of the 1999 IEEE/IEEJ/JSAI International Conference on Intelligent Transportation Systems, October 5-8, 1999, Tokyo, JAPAN.

[21] "Predicting Lane Position for Roadway Departure Prevention" P. Batavia, D. Pomerleau, and C. Thorpe, Proceedings of the IEEE Intelligent Vehicles Symposium, IEEE, October, 1998.

[22] "Model-Based Car Tracking Integrated with a Road-Follower", F. Dellaert, D. Pomerleau, and C. Thorpe,
International Conference on Robotics and Automation, May, 1998.

[23] "A Trinocular Stereo System for Highway Obstacle Detection", T. Williamson and C. Thorpe
1999 International Conference on Robotics and Automation (ICRA '99).

[24] "The 1997 Automated Highway Free Agent Demonstration", C. Thorpe, T. Jochem, and D. Pomerleau
IEEE Conference on Intelligent Transportation Systems, November, 1997, pp. 496 - 501.

[25] "A model driven approach for real time road recognition", R. Aufrère, R. Chapuis, and F. Chausse, International Journal of Machine Vision and Applications, December 2001, Vol. 13, p. 93-107.

[26] "Parallel and Local Feature Extraction : a real-time approach to road boundary detection", A. Broggi, IEEE Transactions on Image Processing, Vol. 4, N°2, p. 217-223, 1995.

[27] "Recursive 3D Road and Relative Ego-State Recognition", E. D. Dickmanns and B. D. Mysliwetz, IEEE Transactions Pattern Analysis and Machine Intelligence, Vol. 14, N°2, p. 199-213, 1992.

[28] "Side Collision Warning Systems for Transit Buses", C. Mertz, S. McNeil, and C. Thorpe, IV 2000, IEEE Intelligent Vehicle Symposium, October, 2000.

[29] "A Performance Specification for Transit Bus Side Collision Warning System", S. McNeil, D. Duggins, C. Mertz, A. Suppe, and C. Thorpe, ITS2002, proceedings of 9th World Congress on Intelligent Transport Systems, October, 2002.

[30]"Driving in Traffic: Short-Range Sensing for Urban Collision Avoidance", C. Thorpe, D. Duggins, J. Gowdy, R. MacLachlan, C. Mertz, M. Siegel, A. Suppe, C. Wang, and T. Yata, in Proc. of SPIE: Unmanned Ground Vehicle Technology IV, Vol. 4715, April, 2002.

[31] "Dependable Perception for Robots", C. Thorpe, O. Clatz, D. Duggins, J. Gowdy, R. MacLachlan, J.R. Miller, C. Mertz, M. Siegel, C. Wang, and T. Yata, Proceedings of International Advanced Robotics Program IEEE, Robotics and Automation Society, Seoul, Korea, May, 2001.

[32] "Environment Recognition from a Car Using an Omni-camera", T. Yata and C. Thorpe, Proceedings of the 19th Annual Conference of the Robotics Society of Japan, September 2001, pp. 539-540.

[33] "Static Environment Recognition Using Omni-camera from a Moving Vehicle", T. Yata, C. Thorpe, and F. Dellaert, tech. report CMU-RI-TR-02-12, Robotics Institute, Carnegie Mellon University, 2002.

[34] "Eye-safe Laser Line Striper for Outside Use", C. Mertz, J. Kozar, J.R. Miller, and C. Thorpe, in IV 2002, IEEE Intelligent Vehicle Symposium, June, 2000.

[35] "Simultaneous Localization and Mapping with Detection and Tracking of Moving Objects", C. Wang and C. Thorpe, in Proceedings IEEE International Conference on Robotics and Automation, May 2002.