Editor: Alberto Broggi
University of Parma, Italy
broggi@ce.unipr.it

Vision in and out of Vehicles

Luke Fletcher, Nicholas Apostoloff, Lars Petersson, and Alexander Zelinsky,
Australian National University

With its wide roads and low congestion, Canberra, Australia, has relatively few serious accidents. However, two recently occurred. In the first, a pregnant woman died in her car while waiting at an intersection.

Her baby daughter was delivered seven weeks early by emergency Caesarian section and is in critical condition. The girl's 10-month-old brother, who was in a child restraint in the backseat, is also in critical condition. The accident happened when a car, thought to have been traveling at up to 80 kmph (the speed limit), failed to stop. Days later, another fatality occurred—this time a truck driver. The truck left its lane, jumped the guard rail, and hit an embankment. Both accidents happened in fine driving conditions on the city's edge.

Perhaps the drivers could have avoided these accidents if they had received a warning about the impending situation. Almost every driver has experienced a warning from a passenger, perhaps about an obscured car while merging or a jaywalking pedestrian in a blind spot. Such warnings could save countless lives every day.

We believe that, in the near future, such assistance will come from the vehicle itself. Many vehicles already employ computer-based driver assistance in the form of antilock braking systems or adaptive engine management systems. However, more than a decade after autonomous-system technologies emerged, systems such as those that Universität der Bundeswehr Munich[1] and Carnegie Mellon University's NavLab group[2] developed have not been realized commercially. A key distinction between existing systems and promising R&D systems is driver and vehicle manufacturer acceptance. Because a paradigm shift to autonomous vehicles is unlikely, intelligent-vehicle technologies must make it to the road by subsystems instead. These subsystems should solve a small, well-defined task that supports, not replaces, the driver.

At the Australian National University's Intelligent Vehicle Project, we are developing such subsystems for

- Driver fatigue or inattention detection
- Pedestrian spotting
- Blind-spot checking and merging assistance to validate whether sufficient clearance exists between cars
- Driver feedback for lane keeping
- Computer-augmented vision (that is, lane boundary or vehicle highlighting on a head-up display)
- Traffic sign detection and recognition
- Human factors research aids

Systems that perform such supporting tasks are generally called *driver assistance systems*. Although we can't know the actual circumstances of the accidents we mentioned earlier, we believe that implementing DAS could prevent similar accidents or at least reduce their severity.

Editor's Perspective

Although road congestion and accidents in Australia are relatively small problems compared to other parts of the world, many laboratories there are developing driver aids. This article describes research activities at the Australian National University. There, a team is studying and testing automatic lane detction, obstacle detction, and driver monitoring on an experimental vehicle.

Feel free to contact me regarding this department. I also seek contributions on the status of ITS projects worldwide as well as ideas on and trends in future transportation systems. Contact me at broggi@ce.unipr.it; ww.ce.unipr.it/broggi. —*Alberto Broggi*

DAS goals

Robustness is of paramount importance for systems in cars driven on public roads. Solutions to sensing and detection problems must be reliable. Fortunately, roads are designed to be high contrast, predictable in layout, and governed by simple rules. This makes the sensing problem somewhat easier, although by no means trivial. Complementary sensors and algorithms can reduce a catastrophic failure's likelihood, but robust systems must incorporate performance metrics and graceful failure modes from the start. Systems must be operable in all driving environments.

This means systems must be able to handle urban environments as well as highways. Urban environments have

1094-7167/03/$17.00 © 2003 IEEE
Published by the IEEE Computer Society

proved difficult owing to an explosion in the diversity of road scenarios and the lower signal-to-noise ratio of content in cluttered urban scenes. Human drivers rely much more extensively on predicting the behavior of other road users and pedestrians in these scenarios than in highway situations. These powers of higher reasoning, which often involve making eye contact with other road users, are not easily modeled and will not come easily to AI systems.

As with most computer vision problems, DAS systems that work 80 to 90 percent of the time are orders of magnitude simpler to implement than systems that work 96 to 99 percent of the time. System designers often face choosing a pedantic system that continually reports false positives or an underconfident system that must continually admit that it can't help at the moment. The ANU Intelligent Vehicle Project addresses this dilemma by observing the driver. Monitoring the driver—particularly where the driver is looking—can avoid many interruptions. That is, if the driver is looking at a potential problem or an uncertain area in the road scene, a warning is irrelevant. This strategy is based on DAS's higher goal to assist drivers by informing them of occurrences of which they might not be aware, not second-guessing drivers' choices when they are paying attention. So, in a complex traffic scene, as long as the driver has noted an identified hazard, such as an overtaking car, or an unpredictable hazard, such as wandering pedestrian, the system gives no alert.

Finally, driver assistance systems must also be

- *Intuitive.* Their behavior makes immediate sense in the context of the standard driving task.
- *Nonintrusive.* They don't distract or disrupt the driver unless they deem it necessary.
- *Overridable.* The driver has ultimate control and can refuse assistance.

Figure 1. The Transport Research Experimental Vehicle uses two vision platforms: the CeDAR (Cable Drive Active Vision Robot) active-vision head and faceLAB passive stereo cameras.

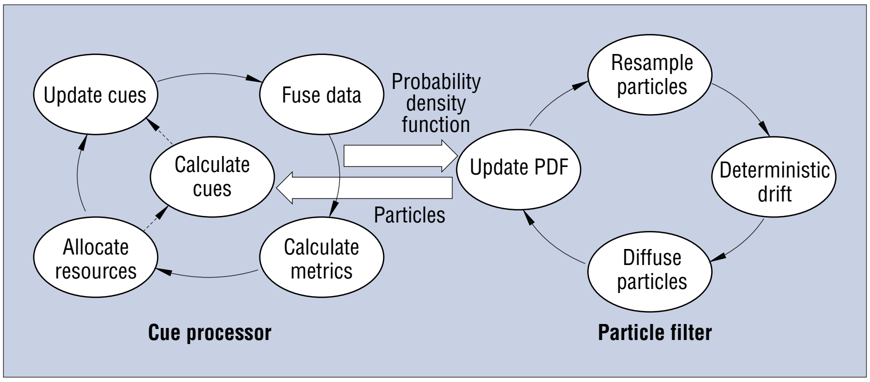


Figure 2. The Distillation visual-cue-processing framework. On the left, visual cues execute on the basis of merit and available computational resources; on the right, a particle filter combines the results.

The Transport Research Experimental Vehicle

The Intelligent Vehicle Project's platform is the *Transport Research Experimental Vehicle,* a Toyota Land Cruiser with a variety of sensors and actuators that support various ITS-related research. Vision is the primary sense used on board TREV, which incorporates two major systems (see Figure 1). An ANU-developed CeDAR (Cable Drive Active Vision Robot) stereo active camera platform, which replaces the rear-view mirror, monitors the road scene in front of the vehicle. This system lets the cameras rotate left and right independently on a shared tilt axis. To monitor the driver, TREV uses a dashboard-mounted faceLAB head-and-eye-tracking system (www.seeingmachines.com/technology/faceLAB.htm).

TREV also includes the typical range of vehicle-monitoring devices: Global Positioning System technology, inertial-navigation sensing, and speed and steering-angle sensors. Throttle and steering actuators support lane-keeping and automatic-cruise-control-style experiments.

Robustness through intelligent use of visual information

Despite many impressive past results in visual processing, no single visual-processing method can perform reliably in all traffic situations. Achieving stable vision-based subsystems will require executing multiple image-processing methods and selecting them on the basis of the prevailing conditions. We have developed Distillation, a visual-cue-processing framework that lets us deal with such robustness issues. Distillation combines a visual-cue-scheduling and data fusion system with a *particle filter* (see Figure 2). (Particle filtering is also

Figure 3. Lane tracker results. Yellow lines indicate the estimated lane boundary. Tracking is robust against shadows, obstacles, and misleading lines on the road.

known as the condensation algorithm or Monte Carlo sampling algorithm.)

Distillation lets us

- Combine visual cues on the basis of Bayesian theory
- Allocate computational resources over the suite of visual cues on the basis of merit
- Employ top-down hypothesis testing instead of reconstructive techniques that are often poorly conditioned
- Integrate visual-cue performance metrics so that we can safely deal with deteriorating performance
- Combine visual cues running at different frame rates

We have also demonstrated Distillation for indoor people tracking.[3]

As part of the Intelligent Vehicle Project, we have developed two applications that exploit the Distillation framework: one for lane tracking and one for obstacle detection and tracking.

Lane tracking

This application combines visual cues such as edges for finding lane marks and road color consistency with cues based on physical-world constraints such as vanishing points and plausible road shapes. From these cues, it distills a winning hypothesis of the vehicle's position with respect to the road and the geometry of the road ahead.

The Distillation framework reduces the lane tracker's search space. Distillation concurrently estimates the road width, the vehicle's lateral offset from the road's centerline, and the vehicle's yaw with respect to the centerline. The application estimates the horizontal and vertical road curvature in the far field.

Figure 3 shows the lane tracker's output in several situations using four different cues.

Obstacle detection and tracking

This application has three main levels. The most primitive level uses a set of "bottom up" whole-image techniques to search the image space for likely obstacle candidates. This level primarily uses stereo disparity and optical flow. Although the disparity and flow information are very noisy, we can combine them to form a 3D depth flow field (see Figure 4). We can then use this field to coarsely segment for potential obstacles using clustering of 3D-flow vectors in the near to medium range.

We can also use color consistency to derive possible obstacle candidates.

The application injects sets of particles representing each obstacle candidate into the particle filter state space inside the Distillation framework. Distillation then tracks the obstacles between frames.

Robustness through driver monitoring

Driver monitoring can

- Reduce false alarm rates.
- Directly detect driver fatigue through head slumping and prolonged eye closure.[4]
- Detect driver inattention to the road scene in general or particular regions.
- Direct DAS attention: "The driver is looking in a particular direction; why?"

If we augment driver monitoring with information about the vehicle and traffic state, we can make additional inferences about the driver. For example, by correlating automated lane tracking and driver gaze monitoring, we can eliminate considerable tediousness from human-factors style experiments.

Seeing Machines developed the faceLAB driver-monitoring system in conjunction with ANU and Volvo Technology Corporation. FaceLAB uses a passive pair of cameras to capture video images of the driver's head. It then processes these images in real time to determine the 3D pose of the person's face (± 1 mm, ± 1 degree) as well as the eye gaze direction (± 3 degrees), blink rates, and eye closure. Figure 5 shows the experimental setup.

The results in Figure 6 show a clear directional bias based on the prevailing road curvature. That is, when driving on a road that curves to the right, the driver focuses more on objects on that side, and vice versa. Figure 7 shows a strong correlation between the vehicle yaw angle from the lane tracker and the gaze direction. That is, the driver is more likely to veer in the direction is which he or she is gazing.

These preliminary results seem to indicate that some of the assumptions about driver gaze that we encapsulated into our subsystems seem justifiable. (For example, two common assumptions are that drivers tend to gaze at the inner edge of curves in the road ahead and that drivers

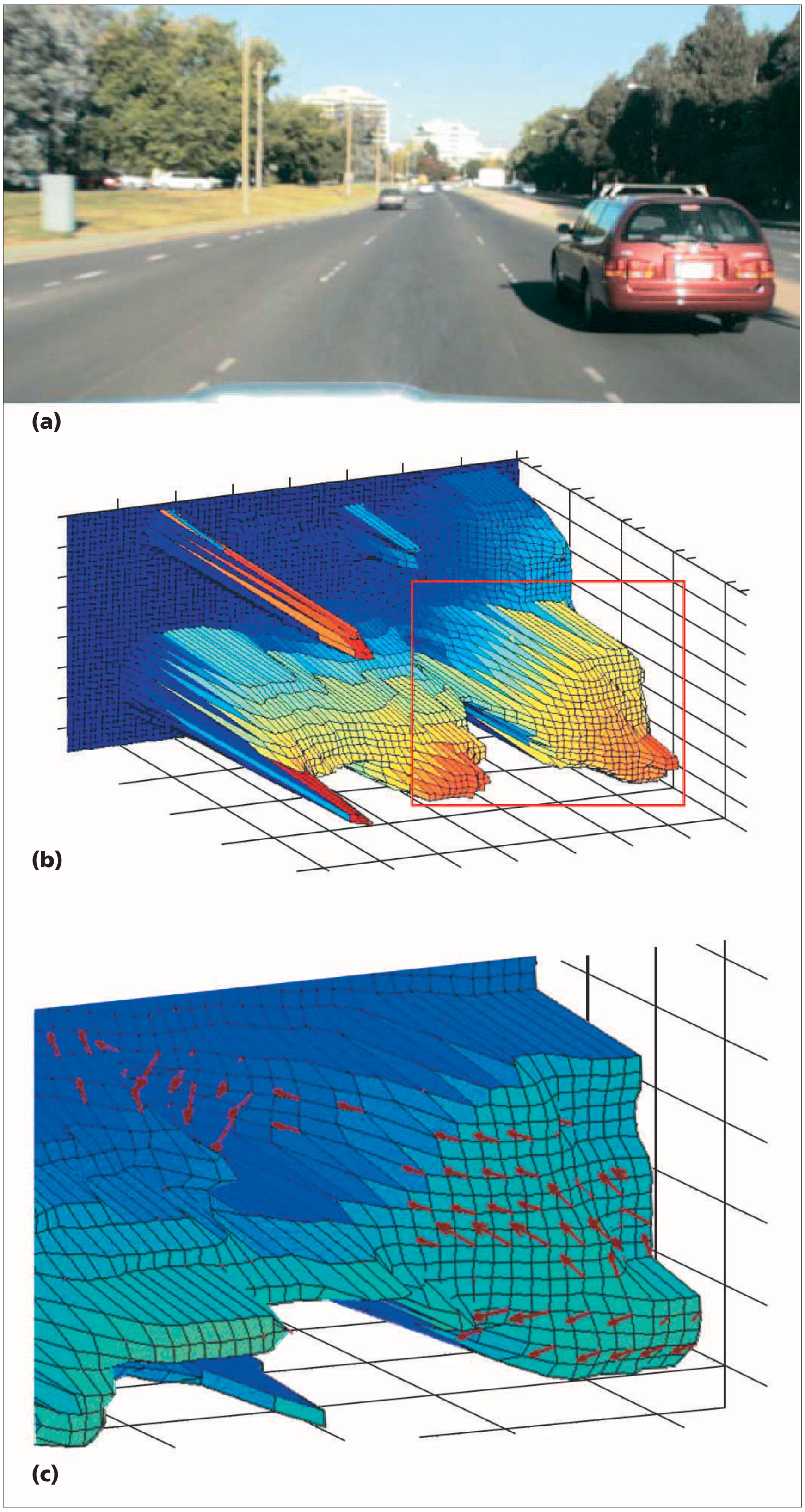


Figure 4. Obstacle detection and tracking: (a) the left image from a stereo pair; (b) a 3D surface from stereo disparity (the rectangle indicates the region of 3D depth flow); (c) 3D depth flow.

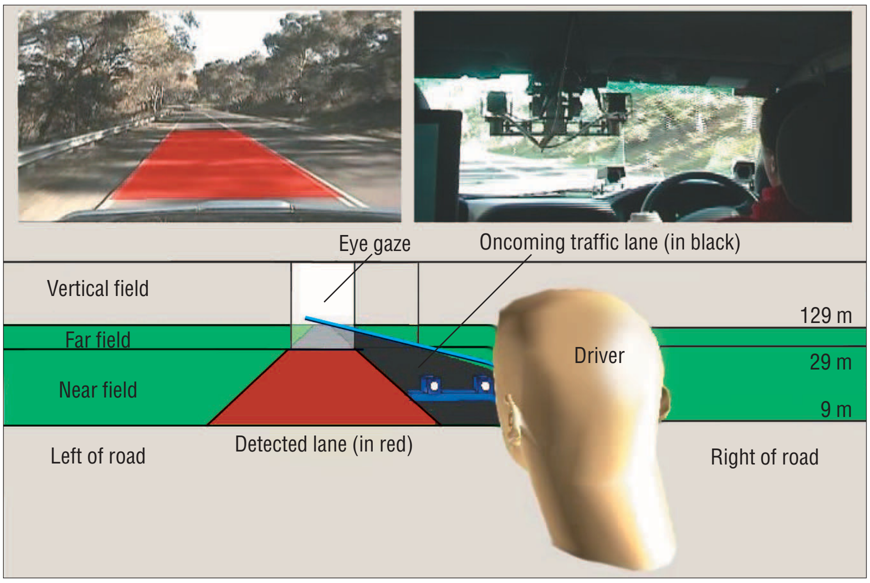


Figure 5. The integration of lane tracking with driver eye gaze tracking. The top left shows the lane tracker's output. The top right shows the video camera arrangement. The bottom shows segmentation of the field of view for analysis.

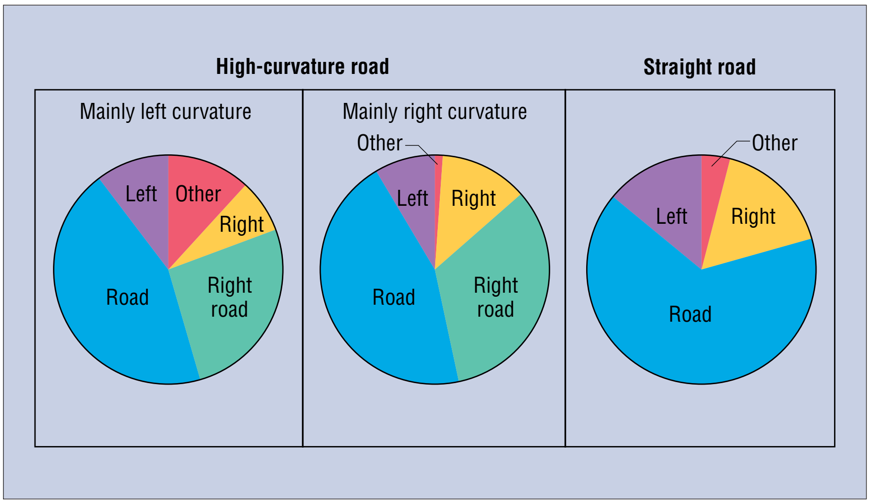


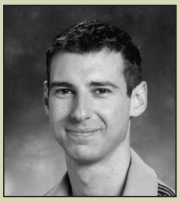
Figure 6. Proportions of viewing direction on left and right curvature roads. When driving on a road that curves to the right, the driver focuses more on objects on that side, and vice versa.

periodically glance [not stare] at the lane of oncoming traffic.) One such subsystem under development is a lane-keeping system that applies a restorative force to the steering wheel if the car approaches the lane boundaries. If the driver looks as if he or she is preparing to change lanes, the subsystem reduces this corrective force in the relevant direction until the lane change is complete.

The Intelligent Vehicle Project is still in its early days. Vehicle and pedestrian detection, blind-spot monitoring, and road-sign detection are all in the pipeline, as are new interfacing systems and context-relevant driver feedback systems such as force feedback and auditory signals.

The need for such driver assistance subsystems is real. Until each vehicle comes

Luke Fletcher is a PhD candidate in engineering at the Robotic Systems Laboratory. Contact him at the Robotic Systems Laboratory, Dept. of Systems Eng., Research School of Information Sciences and Eng., Australian Nat'l Univ., Canberra, ACT 0200 Australia; luke@syseng.anu.edu.au.



Nicholas Apostoloff is a PhD candidate in engineering at the University of Oxford's Visual Geometry Group. He previously was a master's student at the Australian National University researching vision-based lane tracking. Contact him at the Visual Geometry Group, Univ. of Oxford, Oxford, UK; nema@robots.ox.ac.uk.

Lars Petersson is a postdoctoral research fellow at the Robotic Systems Laboratory. Contact him at the Robotic Systems Laboratory, Dept. of Systems Eng., Research School of Information Sciences and Eng., Australian Nat'l Univ., Canberra, ACT 0200 Australia; larsp@syseng.anu.edu.au.

Alexander Zelinsky is a professor of systems engineering in the Australian National University's Research School of Information Sciences and Engineering, where he leads a research team working on intelligent vehicles, mobile robotics, and human–machine interaction. He is also the CEO of Seeing Machines. Contact him at the Robotic Systems Laboratory, Dept. of Systems Eng., Research School of Information Sciences and Eng., Australian Nat'l Univ., Canberra, ACT 0200 Australia; alex@syseng.anu.edu.au.

with its own inbuilt set of vigilant passengers, may the number of lives lost be low. ■

References

1. E.D. Dickmanns and A. Zapp, "Autonomous High Speed Road Vehicle Guidance by Computer Vision," *Automatic Control—World Congress, 1987: Selected Papers from the 10th Triennial World Congress of the Int'l Federation of Automatic Control*, Pergamon, 1987, pp. 221–226.

2. C. Thorpe, Vision and Navigation: *The Carnegie Mellon NavLab*, Kluwer Academic Publishers, 1990.

3. G. Loy et al., "An Adaptive Fusion Architecture for Target Tracking," *Proc. 5th Int'l Conf. Automatic Face and Gesture Recognition*, IEEE CS Press, 2002, pp. 261–266.

4. N.L. Haworth, T.J. Triggs, and E.M. Grey, *Driver Fatigue: Concepts, Measurement and Crash Countermeasures*, Federal Office of Road Safety Contract Report 72, Human Factors Group, Dept. of Psychology, Monash Univ., 1988.

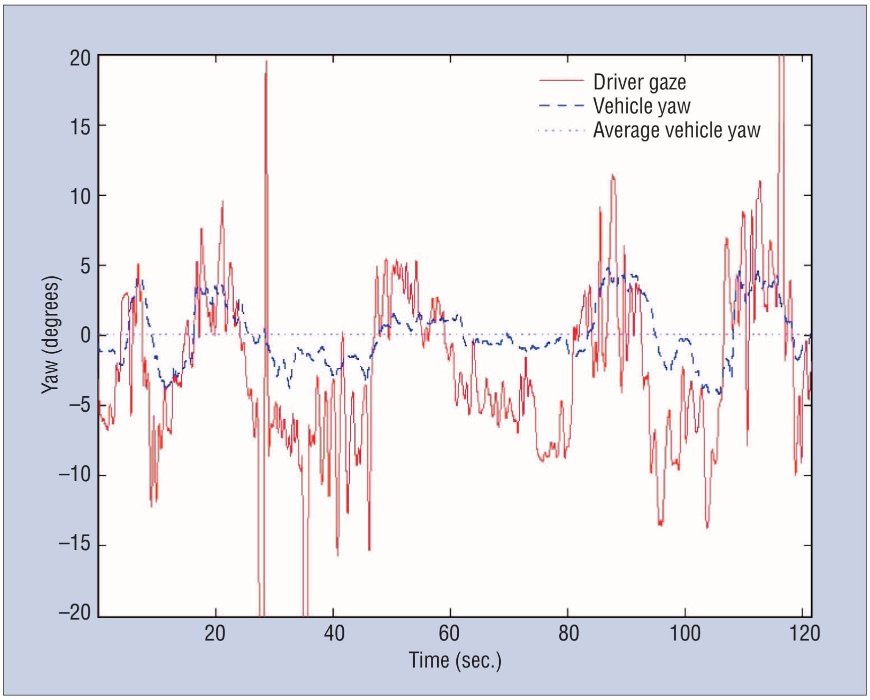


Figure 7. Reconciliation between gaze direction and vehicle yaw angle. The driver is more likely to veer in the direction in which he or she is gazing.