

Machine-vision technology can track road hazards, driver condition, and passenger size, but implementation challenges remain.



ILLUSTRATION BY ELLEN WEINSTEIN

VISION ON WHEELS

By Sridhar Lakshmanan,
University of Michigan, Dearborn,
and Bing Ma and Hyungsoo Kim,
M-Vision Inc.

From the first time they slide in behind the wheel of a car, drivers are told: “Keep your eyes on the road.” Now photonics is improving automotive safety and traffic efficiency with cameras that keep an eye not only on the road ahead but also on the driver and passengers.

A number of technologies are being designed into intelligent vehicles (IV) and intelligent transportation systems (ITS) with an eye to safety and efficiency. Of particular importance among these technologies are those that use photonic sensors. Vision cameras have formed the core of the most-studied and most-widely deployed sensors in IV and ITS due to their fast data-acquisition rate, excellent spatial resolution, wide availability, and promise of low manufacturing cost.

The traditional distinction between IV and ITS is that IV technologies are on the vehicle and benefit the vehicle’s occupants exclusively, whereas ITS encompasses technologies that are part of the highway and provide a broader benefit to all traffic. Road-monitoring IV systems acquire information about the road conditions ahead of and around a vehicle and process such information to understand the environment outside of the vehicle, such as where the vehicle is relative to the edge of a road or a lane, or the precise locations of obstacles in the vehicle’s path. This improved understanding of vehicle surroundings can assist the driver in a number of ways, warning of unintended lane departures, helping avoid collision with obstacles, helping maintain a safe distance with vehicles ahead, and ultimately, perhaps, even assuming complete control of vehicle operation (see table 1).

In contrast, occupant-monitoring IV systems acquire information about vehicle passengers, then process the information to understand the inside environment of a vehicle, including the status of the driver, type of passengers, or number of occupants. With such information, the system can provide

warning to an inattentive or fatigued driver, deploy the passenger-side airbags based on the size and position of the occupant, report the number of vehicle occupants to emergency responders in the event of crash, and so on.

Table 1—Road-monitoring systems

System	Sensor	Processor
GOLD (University of Parma)	stereo CMOS cameras 10 fps output	SIMD parallel computer; (140 ms per frame)
Robert Bosch GmbH	single 512 × 256 pixel 160 fps progressive scan CMOS camera	MMX Pentium with FPGA coprocessor
LOIS (University of Michigan)	single 512 × 480 pixel CCD camera	Pentium board computer (125 ms per frame)
LaneTracker (M-Vision, Inc.)	single CCD camera 30 fps output	Pentium processor (30 ms per frame)
Mitsubishi	single 510 × 492 pixel CCD camera	
SafeTRAC (AssistWare)	CCD camera	AMD DSP
VAMP (UBM)	CCD camera	16- and 32-bit and microprocessor arrays
AutoVue (Iteris)	custom CMOS camera	custom DSP
Mobileye	up to two 640 × 480 pixel CCD cameras	up to two 32-bit RISC ARM946E CPUs, four vision computing engines

Table 2—Occupant-monitoring

System	Sensor	Processor
PERCLOS (Carnegie Mellon)	stereo CCD cameras	PC 104 microcomputer
Eye-tracking (Applied Science Laboratories)	color CCD camera system accuracy: 0.5° system resolution: 0.25°	Pentium type desktop processor
FaceLab (SeeingMachines)	stereo CMOS cameras 60 fps output	Pentium-type desktop processor (200 ms per frame)
OccSens (M-Vision, Inc.)	single CCD camera 30 fps output	Pentium processor (30 ms per frame)
LAAS (CNRS and Siemens Automotive)	stereo 768 × 592 pixel cameras (CCD/CMOS)	Ultra Sparc 10 (250 ms per frame)
Sandia National Laboratories	B&W 480 × 512 pixel CCD camera	
Hughes Research Labs	stereo CMOS cameras 20 fps	400 MHz Pentium II (200 ms per frame)
City University, London University of Applied Sciences, Luneburg	monocular CMOS camera 100 fps	
Daimler Chrysler and Temic	3-D camera 20 fps	

IV in the car

Major automobile manufacturers such as Daimler Chrysler, GM, Ford, Toyota, Honda, Nissan, Volkswagen, and BMW, as well as first-tier suppliers like TRW, Bosch, Valeo, Delphi, Visteon, and Denso, are working on vision-based solutions to render future vehicles safer and more comfortable. Their activity, combined with the progress made at various universities, research institutions, government agencies, and a handful of small- to medium-sized companies, accounts for the bulk of IV technologies.

A typical IV system includes a vision sensor consisting of a focal plane array (FPA) detector, read-out electronics, camera controls, a lens, lens controls, an image acquisition device, an image processor, and an output device or vehicle controller. What distinguishes one IV system from the next is in the performance specifications of these various subsystems: the dynamic range and low-light performance of the detector, the lens shutter speed, acquisition throughput, processor complexity and memory, and control system latency.

IV systems for road monitoring can be divided into two categories. One type encompasses those that use in-vehicle machine vision to provide some sort of automatic detection and tracking of factors such as the shape of the road ahead and the location of other vehicles, pedestrians, traffic cones, and debris that may be ahead, to the side, or behind. The other systems are those that use in-vehicle cameras to provide some sort of enhanced vision such as a view of the blind spot during lane change or backup, a view of the rear for parking assistance, or a 360° mosaic for general situational awareness.

hardware challenges

To monitor the road ahead of and around a vehicle, vision sensors tend to be mounted at a tall location—on the windshield glass, below the interior rearview mirror, on top of the B-pillars (the support between the front and rear side windows), or near the rear brake lights. One of the biggest technical challenges facing vision-based road monitoring is the dynamic range of the detector FPA. On outdoor roads, lighting conditions can vary dramatically due to factors such as time of day, orientation relative to the sun, and weather. In contrast, indoor roads tend to be poorly lit. When the field-of-view includes indoor and outdoor portions, such as at the entry/exit of a tunnel, some areas of the detector array may receive a lot of light, while others receive hardly any.

Conventional wisdom is that CCD detectors seem to provide better dynamic range and noise performance on the low luminance side, whereas complementary metal oxide semiconductor (CMOS) detectors seem to perform better on the high luminance side. The reality is that since the luminance operating conditions are unknown a priori and can vary by an order of 10⁵, both CCD and CMOS FPA detectors are used.

The other major technical challenge facing vision-based road monitoring is image processing. Image-processing algorithms for roadway monitoring are very complex and require a lot of computational resources and memory to keep up with 50-MB/s video rates. Algorithm performance is contingent upon two principal factors: size of the image array and amount of contrast in the image. Ideally, one would like an image with a large number of pixels and very good contrast. Processing a large number of pixels requires considerable

computational resources, however, and contrast can never be guaranteed, so the reality is that most image-processing algorithms are designed with modest assumptions on array size (160×120) and contrast (5:1). The hope is that advances in sensor and processor technologies will allow designers to use the same algorithms to obtain better performance.

A good image-processing algorithm almost always requires a good mathematical model of the spatial relationship between the image pixels. In a complex and widely varying environment such as an outdoor scene, there are no universally good models. Even a simple object boundary—something very apparent to the human eye—is difficult to model due to the unpredictable variability in contrast across the boundary and the variability in shape, orientation, and size of the object. Indeed, the design of robust image-processing algorithms is a major challenge.

vision for occupant monitoring

Compared to road monitoring, occupant monitoring is a more recent area of interest to the IV community. The systems in this field can be grouped into two categories: those that monitor the driver and those that monitor the passengers. The primary reason to monitor the driver is for safety. A study by the U.S. Department of Transportation (USDOT) revealed that roughly one-third of all fatal accidents in the United States are due to drowsy drivers. These are accidents in which a single automobile runs off the road and crashes without involving any other automobile. Another USDOT study concluded that driver inattention was a major contributor to automobile crashes. This inattention problem is likely to grow as more and more telematic devices, which send remote information to the driver, make their way into the automobile. There is tremendous potential payoff, therefore, in monitoring the state of the driver and providing alerts when the driver is inattentive.

In contrast, the primary reason to monitor passengers is to reduce post-crash trauma. Take the case of passenger-side airbags. A recent USDOT study showed that in certain situations airbags could actually be the cause of injury and that the

absence of an airbag might have saved the passenger's life. To prevent such accidental death and injury, efforts are underway to classify occupants by type (adult, child, or infant) and position (seated normally, leaning toward the airbag, etc.). Occupant classification is a fundamental enabling technology for occupant-monitoring systems. It is a tough problem, unfortunately, because there is so little to distinguish between occupant classes. A driver looking at the road in many instances resembles one looking at the instrument panel. A big child in a forward-facing seat, or an infant in a rear-facing seat, can resemble a normally seated small adult.

This lack of clear distinction between the classes holds true regardless of the sensor used to measure occupants, and certainly in data acquired by a video camera. Video images of the driver or the passenger can confuse even human vision, especially when the classification is performed out of context. From a safety standpoint, there's little risk in classifying an attentive driver as inattentive or an adult as a child, while a reverse classification error could be devastating, so there's also the issue of asymmetry between occupant classes.

As one can see, vision is at the core of ITS and IV. That being said, the introduction of vision-based systems by the automobile OEMs is still a few years away. Several orders of improvement in sensor performance and algorithm robustness are needed before these systems are offered to customers as part of a auto dealer package. In the interim, large field operations tests to gauge customer acceptance are being undertaken by the USDOT and its counterparts in other countries. Perhaps a day will yet come where there will be an honest substitute for a real human driver. **oe**

Sridhar Lakshmanan is an associate professor of electrical and computer engineering at the University of Michigan, Dearborn. Bing Ma and Hyungsoo Kim are research scientists at M-Vision Inc., Belleville, MI. For questions about this article, contact Lakshmanan by phone: 313-593-5516; fax: 313-593-9967; or e-mail: lakshman@umich.edu.

Have a comment on this article? Go to the [oemagazine](http://spie.org/app/forums/oemagazine) discussion forum at spie.org/app/forums/oemagazine.