The Intelligent Transportation Society of America (ITS America)

Connected Vehicle Insights
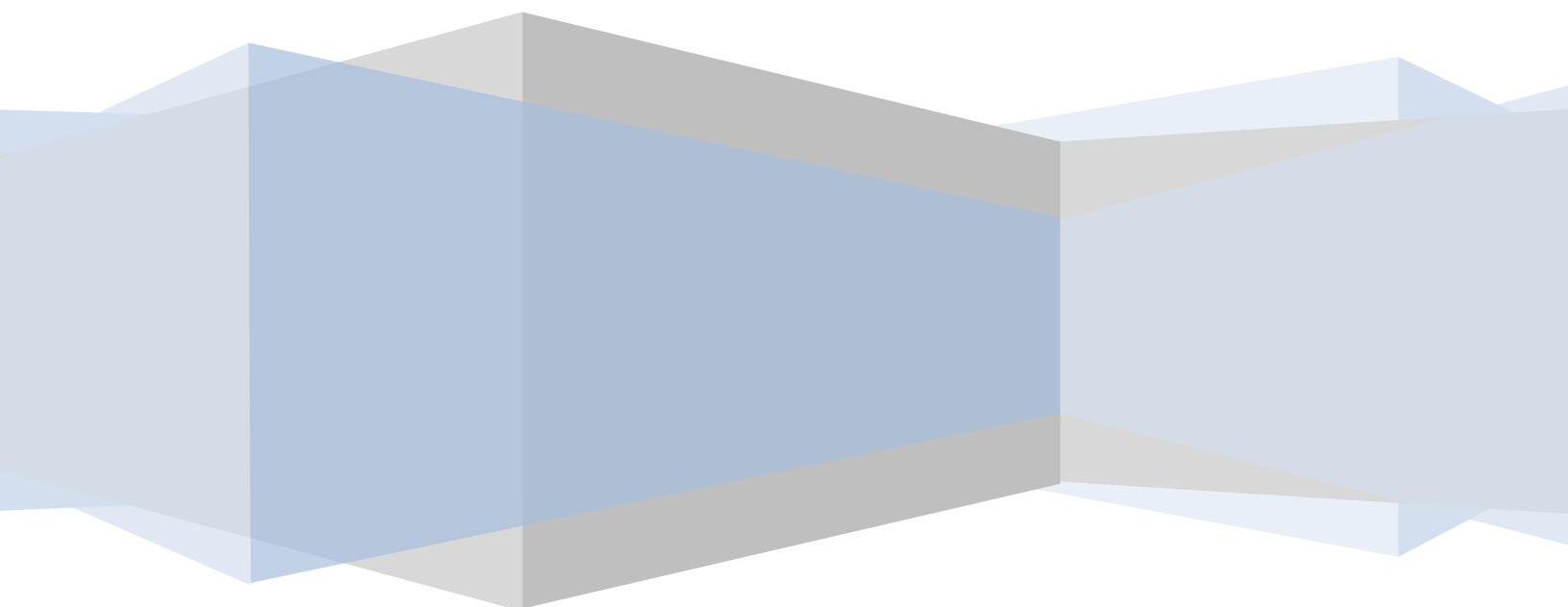
Trends in Computer Vision

An overview of vision-based data acquisition and processing technology and its potential for the transportation sector

Adrian Guan, Steven H. Bayless, and Radha Neelakantan

ITS America Research

Technology Scan Series 2011-2012

Introduction

Drivers must keep their eyes on the road, but can always use some assistance in maintaining their awareness and directing their attention to potential emerging hazards. In the last decade, the auto industry and the auto aftermarket have experimented with devices that provide drivers with a second pair of "electronic eyes," enabled by simple vision-based data acquisition and processing technology.

In 2000, Iteris introduced one of the first commercially available large-scale computer vision applications, *lane departure warning*, in Mercedes Actros trucks.[1] Since then, a number of computer-based vision products have been made available in vehicles and, just recently, in aftermarket automotive devices. By contrast, road operators have for a long time used computer vision to monitor and analyze the performance of their highway networks.

Computer vision is the process of using an image sensor to capture images, then using a computer processor to analyze these images to extract information of interest. A simple computer vision system can indicate the physical presence of objects within view by identifying simple visual attributes such as shape, size, or color of an object. More sophisticated computer vision systems may establish not only the presence of an object, but can increasingly identify (or classify) the object based upon the requirements of an application. In intelligent transportation systems (ITS), computer vision technology is broadly applied to either 1) detect objects and events that may represent safety risks to drivers, or 2) detect hindrances to mobility or otherwise improve the efficiency of road networks.

Computer vision's advantages over many other detection sensors or location technologies are generally twofold. First, computer vision systems are relatively inexpensive and can be easily installed on a vehicle or road infrastructure element, and they can detect and identify objects without the need for complementary companion equipment such as transponders. Second, computer vision systems can capture a tremendous wealth of visual information over wide areas, often beyond the longitudinal and peripheral range of other sensors such as active radar. Through continual innovations in computer vision processing algorithms, this wealth of visual data can be exploited to identify more subtle changes and distinctions between objects, enabling a wide array of ever more sophisticated applications.

The ability to upgrade computer vision processing algorithms allows image sensors, with some limitations, to be improved or even repurposed to support new applications. In vehicles, this means computer vision-based vehicle safety systems may be updated to detect and identify not just other vehicles, but also road signs and traffic signals. For roadside infrastructure, closed circuit TV cameras at an intersection may be repurposed to support not just signal timing, for example, but also traffic incident detection.

Although computer vision can acquire tremendous amounts of visual data, turning it into useful information is in many cases very challenging. Current computer vision technology suffers from the problem of robustness, or the inability to detect objects under a wide variety of operational settings and changing environmental conditions such as illumination and weather. This limitation can be overcome by fusing and integrating computer vision information with data from other external sensors, in a process known as "sensor fusion."

---

[1] "Iteris' Lane Departure Warning System Now Available on Mercedes Trucks in Europe," *IV Source*, June 23, 2000, 1. http://ivsource.net/archivep/2000/jun/a000623_iteris.pdf (accessed March 18, 2011).

Core Technologies and Techniques behind Computer Vision

Two core sensor technologies are the foundation for nearly all computer vision systems. Complementary Metal-Oxide-Semiconductor (CMOS) and Charge-Coupled Devices (CCD) are used on most camera-enabled devices. CMOS is noted for its better performance, battery life, and heat dissipation, thus is mostly used in portable electronic devices such as cell phones. CCD has a similar price point to CMOS and usually has better sensitivity and image quality, and is thus used in most other cameras. Computer vision-based ITS products are not constrained by battery performance, and the difference between the two sensors on other performance parameters is not significant enough to affect the choice between CMOS- or CCD-based image sensors for most transportation applications.

Once visual data has been acquired by the image sensors, it is processed and analyzed by software to enhance image quality and extract key features for either object detection or identification as required by the application. Visual data can be in the form of a single still image, multiple images, or consecutive image sequences, also known as video. When more than one camera is used, the technique is known as *stereo vision*, with the distance between the camera and the object bringing in three-dimensional (3D) features. Since a full spectrum of information may be analyzed and retrieved from an image, e.g. colors, shapes, patterns, and depths, computer vision systems have long been viewed by technologists as a promising approach for a multi-purpose data acquisition or sensing solution.

Key Advantages and Technical Limitations of Computer Vision

The key advantage of computer vision in transportation applications is its non-intrusiveness. In other words, computer vision systems do not need to have devices embedded, physically printed, or externally attached to the objects targeted for detection. In this way, computer vision ostensibly has operational advantages over radio frequency identifier (RFID) tags, barcodes, and wireless access points, which require additional installation of complementary readers, scanners, and wireless modems, respectively. Furthermore, upgrading image sensors does not impose upgrade costs on others. Converting to new image sensor hardware or image processing algorithms does not require upgrading tags, identifiers, or transponders devices on vehicles. For infrastructure-based image sensors, image capturing devices are very easy to mount, remove, replace, and upgrade without extensive lane closures.

Computer vision technology requires significant computing resources. Data processing and analytics in computer vision systems is usually intensive and requires large amounts of computational resources and memory. For example, a simple camera with 800 x 600 resolution is able to capture more than one megabyte per second without image compression (image compression algorithms require additional computational resources).[2] For many ITS applications, this massive amount of data needs to be processed and analyzed in a fairly timely manner. For applications that are not time sensitive, all this data needs to be stored for post-processing, and it is therefore no surprise that many vision-based systems are usually equipped with significant processing memory and data storage.

---

[2] This is based on the assumption of 8-bit grayscale image sequence with 10 frames per second.

The most significant and fundamental technical limitation to computer vision is its robustness in the face of changing environmental conditions. Many transportation applications operate outdoors and are very susceptible to illumination variation such as shadows or other low lighting conditions. Even though some vision-based transportation applications might be immune to illumination variation — for example, driver's infrared night vision — most transportation applications must ensure the scheme's robustness in the face of a large variety of lighting conditions prior to wide-scale deployment.

A rudimentary definition of robustness for a computer vision system is the ability to extract the visual information of relevance for a specific application task, even when this information is carried by a small subset of the data, and/or is significantly different from an already stored image.[3] For example, in segmenting foreground objects from the background, color and shape are often considered to be the main attributes to compare against an existing stored image, but both color and shape can be highly affected by illumination conditions and viewing angles.

Illumination variation further complicates the design of robust algorithms because of changes in shadows being cast. For example, a functional tracking algorithm for vehicles may fail due to the frequent alternation of direct light and shadow of high-rise buildings in urban downtowns. Illumination variation remains the main obstacle for robust computer vision-based ITS applications to overcome.

Robustness of a computer vision-based system requires a sophisticated understanding of application requirements and the ability to scale and cope in the face of potential variation of a number of environmental variables. The development of complex algorithms for applications can be daunting, and workable solutions often require new approaches. The two main potential alternative approaches focus on fusion of data from other sensors or devices, and innovations in algorithm development falling under the category of "machine learning."

The more significant and most often used approach to overcome the limitation of computer vision robustness is *sensor fusion*. Sensor fusion uses techniques that combine data from many sensors in order to produce multiple inferences. In the example of vehicle detection, a radar or LIDAR may be used calculate the changing distance of a target vehicle independently from the computer vision system. This distance measurement is then fed into the computer vision system and allows it to maintain the classification of the size and shape of the vehicle despite changes in angle and distance, or disturbances such as shadows that may obscure features. This process of establishing multiple inferences from many measurements works in a similar way to how human drivers identify vehicles and judge distances. Drivers see and have expectations of vehicle sizes, and how they change relative to distance and angle of view. Drivers also get feedback from the vehicle on how they are accelerating and turning and can therefore maintain, based on a fusion of data from their own senses and the vehicle, a safe distance to the next vehicle.

Furthermore, vehicles with multiple sensors and communication modules could benefit computer vision systems by expanding beyond sensor fusion to integration between devices. Temperature, location, weather, or traffic conditions may have been captured already by other sensors or communication-enabled devices. These independent devices can provide prior information for computer vision system to adjust their preset parameters

---

[3] Peter Meer, "Robust Techniques for Computer Vision," in *Emerging Topics in Computer Vision*, ed. Gerard Medioni and Sing Bing Kang (Prentice Hall: 2004).

or initiate some tasks. For example, if a road is forecast to be covered with ice, a computer vision safety module may adjust itself to more heavily weigh such factors as the changes in color and reflectiveness of the road surface. By making this adjustment in reaction to the reported presence of snow or ice, the computer vision system may adjust application level settings such as the timing of *forward collision warning* (FCW) to a driver to allow greater headway between vehicles, for instance.

Going beyond sensor fusion, *machine learning* may also contribute to the development of more robust computer vision systems over the long term. Machine learning is a sub-field of artificial intelligence focused on the development of algorithms that allow computers to learn from experience based on using example data, such as from sensors or databases. A successful computer vision-based application design chooses to extract visual features that are less likely to be influenced by shadows or illumination. Large "training sets" to exemplify real world scenes are used to tune the algorithm and adjust the parameters. Companies like Google and Microsoft are attempting to apply machine learning to very simple pattern recognition tasks, such as indexing untagged still images to enable their internet search engines to return accurate image query results. The current wide deployment of digital cameras provides abundant image sources for machine learning to draw upon. Machine learning, however, is a relatively new and uncharted field within statistics and computer science, and much still needs to be done in research and development.

Broad Application of Computer Vision in Transportation

Transportation-related computer vision products are being adopted in the consumer marketplace through three sectors: 1) the automotive industry, 2) transportation industrial and infrastructure entities, and 3) the consumer electronics and automotive aftermarket industries.

Despite its distinct technology limitations, the market potential of computer vision systems is still very significant. As image and video capturing hardware, i.e. digital and video cameras, like other IT hardware, have dropped in price significantly in the past decade, dedicated computer vision-based devices have become attractive options for a number of applications, especially in the most-price sensitive markets such as consumer electronics and the automotive aftermarket. For example, in its deliberations on backup cameras, the National Highway Traffic Safety Administration (NHTSA) estimated that an automotive grade camera may only cost between \$58-\$88 for an embedded system.

For industrial and infrastructure entities, ones that support passenger and freight logistics or manage highway networks, computer vision is being adopted to improve operations, lower cost, and improve productivity and mobility. In a few instances, adoption of computer vision is being boosted directly or indirectly by regulation and promotion by government authorities.

In the past five years, there have been a number of new computer vision-based applications for infrastructure operations. *Automatic license plate recognition* (ALPR) was one of the pioneering ITS applications, and is used in highway tolling as an enforcement backup to tolling transponders. By scanning the images of a license plate, an ALPR application is able to convert license plate characters into machine-encoded text and compare the results against a database of valid vehicles. In the last several years, there has been a lot of experimentation in the area of

video tolling, an ALPR application. Moving beyond tolling enforcement, video tolling may be a cost-effective way to toll, taking advantage of computer vision's non-intrusive quality. Without requiring deployment and management of complementary transponders, video tolling may be another technology alternative for road operators seeking new ways to throttle and manage traffic demand via road pricing.

Computer vision technology also plays many other roles in improving productivity and safety in traffic management and other transportation operations. Many surveillance cameras have been mounted along freeways and main intersections for public safety, traffic incident detection, ramp metering, and traffic signal timing. They have also been used as backup to calibrate or diagnose problems with other traffic sensors such as loop detectors. With the wide deployment of roadside surveillance cameras and computer-aided analysis of traffic surveillance video, transportation planners and researchers have the ability to extract traffic parameters and patterns beyond just speed and flow, such as vehicle lane changing or acceleration patterns. In addition, computer vision systems can help staff in Traffic Management Centers monitor more operational scenes concurrently and act more quickly in response to emergencies such as crashes or other adverse traffic incidents.

Data collection, mapping, and asset management is another broad category of applications that may utilize computer vision to improve productivity for transportation industrial and infrastructure entities. Transportation agencies may drive highway miles with computer vision imagers to identify the location and condition of assets such as traffic control devices and signs, or survey critical mobility and safety characteristics of road networks, such as areas where there are dangerous curves, rough pavement, or potholes. Commercial map databases from NAVTEQ, Tele Atlas or Google in many cases lead the way in this area, with "street view" imagery which adds visual context to traffic, navigation, and points of interest data.

The value proposition for computer vision technologies in the auto industry to date has been focused on safety, through crash avoidance applications. Computer vision-based lane departure warning (LDW) systems have become a feature in light-duty vehicles, in particular in the higher-end offerings of such manufacturers as General Motors, BMW, and Volvo.[4] A vehicle-based lane departure warning system is an embedded or aftermarket device designed to warn a driver when the vehicle begins to move out of its lane on freeways or arterial roads. In 2010, Volvo even introduced a computer vision-based Pedestrian Detection Safety System.[5] By fusing images with data from other sensors, the 2010 Volvo S60 took computer vision commercialization a step further by giving applications the ability to differentiate pedestrians from other objects in the driver's forward field of view.

For a long time, automotive manufacturers have included computer vision-based *advanced driver assistance systems* (ADAS) as optional or standard equipment in their high-end luxury vehicles. However, inexpensive hardware and regulatory incentives may soon push computer vision applications down to lower-end vehicles as well. NHTSA, for example, has begun to improve the New Car Assessment Program (NCAP) to recognize improvements in crash avoidance systems. Known as the NHTSA "stars" rating program, the NCAP was designed to provide consumers with criteria upon which to judge the safety of new cars prior to purchase, in an effort to encourage consumers to purchase safer cars. The upgrade to NCAP "stars on cars" has already begun and includes

---

[4] Mobileye, "Lane Departure Warning." http://www.mobileye.com/node/58, (accessed March 18, 2011).

[5] "2011 Volvo S60 to Include Pedestrian Detection Safety System," *Edmunds Daily*, April 2, 2010. http://blogs-test.edmunds.com/strategies/2010/04/2011-volvo-s60-to-include-pedestrian-detection-safety-system.html (accessed March 18, 2011).

not only star ratings for crashworthiness, but also recognition for crash avoidance systems such as lane departure and forward collision warning.

Market penetration of ADAS applications, many of which rely upon computer vision, will likely increase due to lower component costs, application maturity, and market demand for safety features, in addition to regulatory incentives. For example, lane departure warning systems, some including automatic steering correction, will be available on high-volume models such as the Ford Focus and Mercedes-Benz C-Class by 2011. According to ABI Research, annual lane departure warning installations are projected to reach over 22 million, with a worldwide market value of more than $14.3 billion by 2016.

Although growth in ADAS in light vehicles may pick up, automotive manufacturers of heavy vehicles such as semi-trailers and other trucks have in the past been early adopters of computer vision-based systems. Commercial freight operators and their automotive equipment suppliers have been in the forefront of the development of computer vision-based ADAS for a couple of reasons. First, commercial freight operators, unlike consumers, are incentivized financially and through regulation to reduce their risk exposure and insurance costs by introducing collision avoidance applications to their trucking fleets. Second, trucks are built to order by trucking companies and customization is the rule, not the exception as in light-duty vehicle manufacturing. This allows newer technologies to be added to trucks in a more flexible and practical way, based upon clear assessments of the cost and benefits to the carrier. Since commercial carriers are generally more cognizant of the costs of crashes and the impact to their business' reputation, goodwill, and bottom line, they have been the most aggressive adopters to date of lane departure warning and forward collision warning systems, many of which include computer vision technologies.

In the near future, it is likely that almost every new light vehicle will have at least one computer vision-based system. NHTSA is considering a rule that could potentially mandate that all vehicles sold in the US after 2014 include a rear-mounted camera and an in-vehicle display. This backup camera system would be designed to reduce the nearly 18,000 back-over incidents that occur in the US per year.[6] New backup camera hardware may serve as a platform that could support additional advanced computer vision-based applications such as *intelligent parking assist systems* (IPAS), which automates parallel parking maneuvers.[7]

Finally, there have been new offerings of computer vision products in the consumer electronics or automotive aftermarket that are focused on safety and mobility. Computer vision is well suited to vehicle aftermarket applications because of the image sensor's small form factor. Additionally, image cameras that are mounted on rooftops or windshields are easy to install. Several companies, such as Iteris/Audiovox and Mobileye,[8] have already developed aftermarket standalone or retrofitted products that can be mounted on a car windshield and can distinguish between pedestrians, bicycles, and vehicles for forward collision warning. Some aftermarket

---

[6] US Department of Transportation National Highway Traffic Safety Administration, "Federal Motor Vehicle Safety Standard, Rearview Mirrors; Federal Motor Vehicle Safety Standard, Low-Speed Vehicles Phase-In Reporting Requirements. Notice of Proposed Rulemaking." http://www.gpo.gov/fdsys/pkg/FR-2010-12-07/pdf/2010-30353.pdf (accessed March 18, 2011).

[7] ABI Research, "Press Release: Lane Departure Warning Systems Go Mainstream: $14.3 Billion Market by 2016," February 22, 2011. http://www.abiresearch.com/press/3619-Lane+Departure+Warning+Systems+Go+Mainstream%3A+$14.3+Billion+Market+by+2016 (accessed March 18, 2011).

[8] John Quain, "Lane Departure Warning Systems for the Rest of Us," *New York Times Wheels Blog*, http://wheels.blogs.nytimes.com/2010/04/13/lane-departure-warning-systems-for-the-rest-of-us/ (accessed March 18, 2011).

vehicle-based operator alertness/drowsiness/vigilance monitoring devices use computer vision to detect eye movements, blinking, and head poses to predict drowsiness in drivers.[9]

The wild card for computer vision's adoption in transportation is the development of new portable connected computing platforms such as smartphones and tablets. Most smartphones and tablets (and soon likely other consumer electronic device or appliances, and possibly even new lines of personal navigation devices) already contain one or more embedded cameras and can serve as computing and sensor platforms for vision-based applications. By creating a third-party developer community and web store for applications, smartphones in many ways have already unsettled previously well-established automotive aftermarket product categories, such as GPS or personal navigation devices. As a result, several companies such as Nokia and QNX have been in a rush to integrate these portable phone/tablet computing platforms into the telematics units of vehicles.

Another possibility for computer vision is the intersection between mobile internet devices, "cloud services," and device-embedded imagers. Equipped with navigation (GPS), gyroscope and accelerometer sensors, devices are now beginning to implement mobile "augmented reality" interfaces, overlaying information queried from online "cloud-based" servers about live moving images to help the user quickly analyze and understand their immediate environment. The market research firm ABI estimates that the market for mobile augmented reality products will grow from $21 million today to an astonishing $3 billion dollars by 2016.[10]

Augmented reality interfaces and other data visualization systems incorporating image sensors may be important in data management, mapping, and asset applications in the transportation domain. However, it is not clear how mass market consumers will embrace augmented reality without a better understanding of the potential benefits and limitations.

Specific Market Potential for Computer Vision in Safety Systems

Most current or potential safety-focused ITS applications that depend upon computer vision technology are typically very sensitive to processing delay and must operate in real time. Infrastructure-based applications falling into this category include actuated traffic signals, ramp control, *curve speed warning* (CSW), and *intersection collision warning* (ICW). Vehicle-based applications include *lane departure warning* (LDW) and *lane-keeping assistance* (LKS), *adaptive cruise control* (ACC, also known as headway detection), *forward collision warning* (FCW), *intelligent parking assist systems* (IPAS), night vision enhancement, and object detection and warning for pedestrians, bicyclists, or animals.[11] Others have less stringent latency requirements, e.g. freeway or arterial surveillance, pavement condition, and drowsy driver warning systems.

---

[9] Lawrence Barr, et al. "A Review and Evaluation of Emerging Driver Fatigue Detection Measures and Technologies," US Department of Transportation. http://www.ecse.rpi.edu/~qji/Fatigue/fatigue_report_dot.pdf (accessed May 11, 2011).

[10] ABI Research, "Press Release: Augmented Reality-Enabled Mobile Apps Are Key to AR Growth," February 11, 2011. http://www.abiresearch.com/press/3614-Augmented+Reality-Enabled+Mobile+Apps+Are+Key+to+AR+Growth (accessed March 18, 2011).

[11] US Department of Transportation Research and Innovative Technology Administration, "Intelligent Transportation Systems Application Overview." http://www.itsoverview.its.dot.gov/ (accessed March 18, 2011).

There are likely three major successive trajectories in vehicular safety applications that may rely in whole or in part on computer vision. The first trend is the commercialization of simple, standalone, and single-purpose driver assistance applications for automotive aftermarket systems. The second trend is more complex embedded ADAS applications that can fuse other sensor and vehicle data with computer vision inputs. The third trend, much further in the future, is highly sensor-fused, computationally-intensive complex systems that are robust enough not just to provide warnings to drivers, but to provide automated inputs into vehicle control systems such as braking, steering, and other dynamic controls.

The first trend is commercialization of computer vision applications focused on simple object detection and identification to provide advisory alerts drivers. The major limitation to computer vision, the lack of robustness of image processing algorithms, can be ameliorated most effectively by focusing the vision-based sensor and data processing resources on identifying very simple predictable objects. Lane departure warning is a good example of such narrow application focus, as road surface markings have simple distinct colors, like white or yellow, and geometric patterns, such as straight or curved lines that are easy to predict, detect, and identify. In a similar vein, because color, shape, and symmetry of traffic signs can be exploited for easy detection, dedicated vision-based applications in vehicles may, for example, detect and interpret traffic lights or traffic signs as well to provide advisory messages or warnings to drivers. For example, Ford recently introduced Traffic Sign Recognition to alert drivers when they are driving too fast for the speed limit. Driver alertness monitoring also fits this category of simple detection and alerts based upon easily detectable features (e.g. indications such as peculiar eye or head movements) for identifying fatigue or sleep deprivation.

The second trend is computer vision fusion with vehicle data or other sensors to provide ADAS applications. Since a computer vision system is capable of capturing a limited number of target features effectively, it can be of more use if integrated with other sensing and vehicle control systems that provide improved driving context. A lane departure warning system, for example, activates a warning when the vehicle appears to be drifting into the next lane, but only alerts if the driver has not properly activated his turn signal. Other ADAS applications only work at low or limited speeds, based on the speedometer, to reduce the risk that the computer vision system may not be fast or robust enough to manage all operating conditions without generating false or otherwise imprudent warnings. The *pedestrian collision warning* (PCW) for Mobileye's camera system works at speeds between 1 and 45 mph, and its lane departure warning operates up to 75 mph. Mobileye's forward collision warning systems work at speeds above 30 mph.

Some ADAS are not only integrated with vehicle control systems and driver interfaces, but are also sensor-fused. The recent pedestrian detection system developed for Volvo uses the sensor fusion approach. The Volvo XC60 uses a combination of a radar sensor and a camera to identify standing or moving pedestrians within a 60 degree field of view in daylight. The Volvo XC60 radar detects objects of any shape and measure the distances to them, and the camera determines, based on the contour of the entire object line, whether the object is a pedestrian. With other situation-aware sensors added to vehicles, such as radar, laser, or infrared thermal measuring devices, it is expected that more sensor fusion-based applications might emerge to fully leverage the strengths and balance out the weaknesses of computer vision-based systems.

The third trend is computer vision systems that provide inputs into vehicle control systems such as braking, acceleration, and steering to support ADAS or even *autonomous vehicle systems* (AVS). Most time-sensitive or

safety vision-based applications focus generally on assisting the driver with warnings, rather than using data to directly influence safety-critical vehicle control systems. Vision-based technology that is fused with a number of sensors, as seen in experimental military autonomous vehicles or concept cars such as Google's Driverless Car,[12] are different in that they move beyond driver assistance to the integration of sensors into the control systems of the vehicle powertrain, body, and chassis. These cars use multiple vision and non-vision based sensors, digital maps, actuators, and complex and sophisticated algorithms to fuse data to support multiple complex driving tasks, not just single applications, without driver intervention. These systems feature real-time intelligent controls that adapt to uncertain and unstructured operating environments.

"Connected vehicle" or cooperative safety systems will likely benefit from computer vision technology. Connected vehicle systems are different from ADAS in that they rely on data generated by other vehicles, not just from on-board sensors, devices, or controllers. Connected vehicle applications could fuse sensor data from neighboring vehicles to support cooperative collision avoidance applications, using, for example, dedicated short range communications (DSRC), GPS, and possibly other sensors. DSRC is a critical technology in connected vehicle or cooperative safety systems. DSRC can transmit information such as speed, heading, and other data that can be used by other vehicles to generate warnings or plan safe maneuvers, even at high speeds.

ADAS or AVS may be linked between vehicles using DSRC to coordinate and confirm vehicle movements such as in *emergency electronic brake lights* (EEBL), *adaptive cruise control* (ACC), or *platooning*, three applications that manage and coordinate acceleration and braking to maintain safe longitudinal headways between two or more vehicles. For example, instead of relying solely on GPS, a connected vehicle safety system could use computer vision to detect a braking vehicle ahead, warning the driver. The connected vehicle system utilizing computer vision could then use DSRC to transmit an "electronic brake light" warning to vehicles behind it which would need to prepare for hard braking.

Lastly, "connected vehicle" cooperative safety may include *intersection collision avoidance* (ICA) applications. ICA runs where traffic controllers communicate to vehicles approaching an intersection to advise drivers of potential risks, or to coordinate vehicle movements. Vehicle-intersection communications using DSRC is also an opportunity for computer vision. Of the nearly 250,000 controlled intersections in the US, a growing number use computer vision-based systems to time and actuate traffic controllers. In the future, computer vision technology may be able to identify the speed or maneuvering of vehicles near or inside an intersection to identify potential dangers, such as the risk of running a red light, or a potentially dangerous left turn onto rapidly oncoming traffic. A computer vision system may be able to detect and identify vehicles that are not equipped with DSRC or GPS.

Computer Vision Adoption and Implications for New Vehicle and Infrastructure Technologies

Computer vision technology is ripe for further application development in transportation. Computer vision technology is a natural solution because of its unique properties: non-intrusiveness, flexibility in fusion with other sensor technologies, and ability to support a variety of application with updates and changes in algorithms.

---

[12] Sebastian Thrun, "What We're Driving At," *The Official Google Blog*, http://googleblog.blogspot.com/2010/10/what-were-driving-at.html (accessed March 18, 2011).

Computer vision will likely be a sought-after core or complementary technology for a number of new ITS applications.

There are a number of vehicle and infrastructure applications within the state of ITS practice today where the performance and reliability of computer vision is more than sufficient by itself. Even though driving a car is a highly dynamic activity, not all vision-based on-board applications are required to be conducted in real time to be of value. In road weather detection, rain or fog may be detected in seconds or minutes instead of milliseconds; a drowsy driver alert may still be of use when triggered minutes after the first sign of fatigue is detected. Less time-critical applications, for example image sensors that change the signal phase and timing of traffic controllers based upon the flow and number of vehicles queued at an intersection, have been deployed in large numbers for quite some time. Despite challenges in maintaining the accuracy of automatic number plate recognition in the face of many different license plate designs, serial formats, and vehicle mounting standards, "video tolling" is also gaining in deployment. The number of computer vision applications that reach maturation in the market for ITS products and services is likely to expand at a reasonably steady pace.

There are also application areas in which computer vision performance is improving and in some instances overtaking other technology solutions. In most cases, computer vision systems are being adopted as a complement to existing technologies, such as transponders, radars, detectors, or other location technologies, but over time confidence in the technology has improved to such a degree that it is being deployed by itself (e.g. video tolling at some road facilities, where the performance of computer vision was deemed equal to that of transponders). In this way, computer vision is a potentially "disruptive" technology for a limited set of applications, assuming improvements in robustness and overall application and system reliability and capability. However, for many applications, especially mission-critical and safety-focused applications, computer vision systems are generally not robust enough by themselves, and applications may require additional backup systems such as range-finding (radar, LIDAR) or transponders.

Because computer vision systems can capture a tremendous wealth of visual information over wide areas, the hope of technologists is that such systems can be updated, through improvements in algorithms, to accommodate newer varieties of applications using the same sensor platform. Ultimately, the performance of computer vision-based solutions must be evaluated based on requirements for performance, reliability, and cost, and these requirements may differ considerably across applications. Even though computer vision holds promise as a flexible "general purpose" sensor like GPS, there may be significant limitations in its ability to pool multiple applications from the same imaging platform.

Computer vision systems will likely be part of a sensor suite that supports multiple vehicle safety applications. Although autonomous vehicle technology will not be commercialized in the next few years, ADAS applications such as lane departure warning or blind spot detection might be boosted with the aid of vision-based technology. Integration of communications using DSRC will ensure that sensor information from vision-based systems will be shared not only between applications, but perhaps even between vehicles. Computer vision systems may act as a check on other sensors, ensuring that the total reliability of vehicle safety applications is maintained.

Assuming global vehicle manufacturing is nearly 50 million vehicles per year, it is likely that nearly half of all new vehicles worldwide will include a forward-facing camera to support lane departure warning, and a backward-facing

backup camera, should NHTSA be aggressive in its rulemaking on rearview systems. Computer vision systems may become as common as GPS equipment currently installed by auto manufacturers for telematics systems, or brought in by consumers as smartphones or personal navigation devices.

There is no easy formula for accelerating consumers' adoption of a new technology. Generally, the automotive manufacturing sector lags technology trends from the IT and consumer electronics industry by several years, waiting to gauge technology maturity and user acceptance. However, automotive manufacturers have been keen to accelerate technology in their vehicles as explosive growth has occurred in mobile broadband and computing. The light-duty vehicle industry is beginning to recognize that younger car buyers will soon be picking vehicles not just by their horsepower, styling, or overall practicality, but also by their connectivity and technology. Changes in consumer preferences mean that drivers will be more receptive to and may see the value in computer-based vision applications in cars.

The Next Steps in Computer Vision

Just fifteen years ago, it was difficult for anyone to imagine that nearly every new phone would have a camera and GPS, plus a compact computer operating system that allows third-party developers to create applications that take advantage of these sensors. If computer vision-based systems were installed in nearly every new vehicle and openly accessible to third-party application development in the same way as smartphones or tablet computers, there could be tremendous opportunities for new applications that take advantage of the roaming vehicle- or pedestrian-based camera. For example, the cost of data collection for mapping and asset management may drop if even a small set of commercial vehicles can be set up to act as volunteer "computer vision" probes. Probe data is currently collected using GPS on commercial vehicles equipped with fleet telematics that collect sample traffic speeds for processing and dissemination by the major commercial traffic information service providers. Computer vision probe data could be collected in a similar way to detect and identify changes to road surfaces, signs or other assets and points of interest.

Furthermore, moving beyond the use of computer vision systems for simple object detection in vehicle or roadway applications, the development of augmented reality systems may provide future human-machine interfaces that could influence a number of transportation applications in the automotive, industrial, and infrastructure sectors. Not only will augmented reality interfaces query and present data based on location and orientation, they could conceivably provide some automated object detection and identification functions.

Through augmented reality, a "crowdsourced" data infrastructure could eventually come into being to share computer vision imagery, video, and processing algorithms. Specifically, because of the improved ease-of-use of augmented reality interfaces, these systems may serve as new platforms to gather data, as imagers may be pointed by users to survey and annotate objects of interest to be stored in geographical information systems (GIS). Unlike Google or NAVTEQ street level imagery, however, data for these systems will likely be managed not by GIS professionals, but may be collected by anyone in a crowdsourced or peer-to-peer fashion. This crowdsourced data could be transmitted, stored, and made available online, but would require new scalable ways of processing the enormous amounts of visual data that will be generated.

Computer vision-based object detection and identification algorithms will likely need to be developed by a large and open developer community to ensure the quality and usefulness of annotated crowdsourced data for new applications. If such large-scale crowdsourcing and processing automation succeeds, the visual digitization of the physical world through computer vision systems could have a tremendous impact on transportation. Just as GPS has been a vital tool, computer vision may be of almost equal importance in the future as an instrument used to improve the efficiency, reliability, and safety of our transportation system.

About the *Connected Vehicle Technology Scan Series*

Under sponsorship from the US Department of Transportation (USDOT) Intelligent Transportation Systems Joint Program Office (ITS-JPO), the *Intelligent Transportation Society of America* (ITS America) is conducting the *Connected Vehicle Technology Scan and Assessment* project.

This two year scanning series of *Connected Vehicle Insight* reports will assess emerging, converging, and enabling technologies outside the domain of mainstream transportation research. ITS America seeks technologies that will potentially impact state-of-the-art or state-of-the-practice in ITS deployment over the next decade, with an emphasis on the "connected vehicle."

The *Technology Scan Series* notes trends, technologies, and innovations that could influence, or be leveraged as part of, next-generation intelligent transportation systems within the next five to seven years. The series' focus is on developments in applied science and engineering and innovation in data acquisition, dissemination, processing, and management technologies and techniques that can potentially support transportation.

To learn more about the Technology Scan Series, and to see more *Connected Vehicle Insight* reports, please visit http://www.itsa.org/knowledgecenter/technologyscan.

To learn more about USDOT's research in Intelligent Transportation Systems, please visit http://www.its.dot.gov.