IEEE pervasive COMPUTING
MOBILE AND UBIQUITOUS SYSTEMS
www.computer.org/pervasive

Turn-Intent Analysis Using Body Pose for Intelligent Driver Assistance

Shinko Yuanhsien Cheng and Mohan Manubhai Trivedi

Vol. 5, No. 4
October–December 2006

Turn-Intent Analysis Using Body Pose for Intelligent Driver Assistance

Research for a proposed driver-turn prediction system examines how and to what extent you can use body-pose information to detect driver intent and behaviors.

Shinko Yuanhsien Cheng and Mohan Manubhai Trivedi
University of California, San Diego

Human-centric, pervasive computing environments, with integrated sensing, processing, networking, and displays, provide an appropriate framework for developing effective driver-assistance systems. Also essential when developing such systems are systematic efforts to understand and characterize driver behavior.

For example, imagine a scenario in which a motorist tries to turn right at an intersection and a bicyclist is in his blind spot. If the vehicle could detect the bicyclist in its path, it could alert the driver and potentially avoid the collision. However, if the driver has already seen the bicyclist, an alert might hinder rather than complement safe driving by unnecessarily increasing the driver's workload. Consequently, for an intelligent driver-assistance system to be effective, it must be able to continuously monitor not just the surrounding environment and vehicle state but also the driver's behaviors. That way, if a dangerous situation occurs that requires intervention, the vehicle can recognize it and accurately alert the driver.

In an attempt to make such a predictive turn-assistance safety system a reality, we equipped an experimental vehicle with cameras and sensors to capture the vehicle dynamics, view of the road ahead, and driver's body pose. We investigated how and to what extent we could use body-pose information to detect and predict driver activities. We analyzed the detection performance of a two-class pattern classifier using receiver-operator-characteristic curves, which describe the classifier's ability to suppress missed detections and false alarms. The curves provide a ratio indicating the system's attainable *proactivity* (ability to foresee a user's needs) versus its *transparency* (ability to avoid user annoyance).[1] Our goal is to eventually develop vision-based body-pose-recovery and behavior-recognition algorithms for driver-assistance systems.

Holistic sensing for active safety

There has been considerable work on systems for sensing the environment, but only recently has there been work on systems for sensing the driver and other vehicle occupants (see the "Related Work in In-Vehicle Sensing" sidebar). Such sensing should be very useful for drivers, who are inherently limited in their ability to maintain accurate awareness of the driving context, especially when the situation is complex, the weather adverse, or the task load too high. Driver assistance should provide a rich, informative environment without adding to the driver's workload. In developing systems that proactively assist drivers (and occupants), human intent is a critical piece of information for determining whether the system's actions will help or hinder the user.[1]

As illustrated in the bicyclist example, an alert

Published by the IEEE CS and IEEE ComSoc ■ 1536-1268/06/$20.00 © 2006 IEEE

Related Work in In-Vehicle Sensing

There has been considerable work on systems for sensing the environment, such as forward-pedestrian detection, blind-spot obstacle detection, ahead vehicle-proximity sensors for stop-and-go cruise control, rear-bumper proximity sensor for back-up assistance, rain sensors for automatic windshield wipers, and surround imaging for parking assistance (see figure A).[1–5] Examples of systems for vehicle state sensing include accelerometers for airbag control and emergency door unlocking, GPS for navigation, and on-board sensors for vehicle parameters such as speed, steering-wheel angle, yaw rate, gear position, throttle position, acceleration, and brake pressure.[6,7] Only recently has there been work on systems for sensing of the driver and the vehicle's other occupants. Examples of these systems include a vision-based occupant posture recovery system[8,9] and occupant type classification for full-force, half-force, and no airbag deployment as well as head pose, eye gaze, and eye closure statistics estimation for driver vigilance monitoring and driver view synthesis.[10] There are also some interesting preliminary results in body posture analysis for classification between safe and unsafe postures of drivers.[11,12]

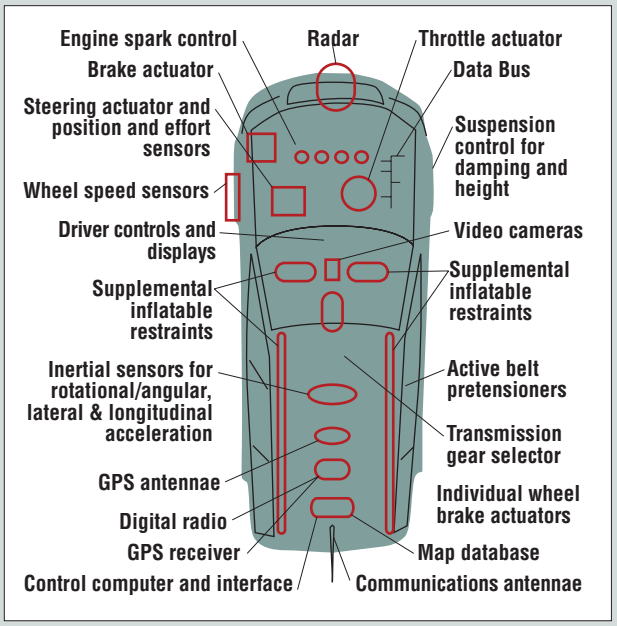


Figure A. A sensing, computing, networking infrastructure for today's vehicles.

REFERENCES

1. Z. Sun, G. Bebis, and R. Miller, "On-Road Vehicle Detection: A Review," *IEEE Trans. Pattern Recognition and Machine Intelligence*, May 2006, pp. 694–711.
2. T. Gandhi and M.M. Trivedi, "Vehicle Surround Capture: Survey of Techniques and a Novel Omni-Video-Based Approach for Dynamic Panoramic Surround Maps," *IEEE Trans. Intelligent Transportation Systems*, Sept. 2006, pp. 293–308.
3. F. Heimes and Hans H. Nagel, "Towards Active Machine-Vision-Based Driver Assistance for Urban Areas," *Int'l J. Computer Vision*, Oct. 2002, pp. 5–34.
4. L. Li et al., "IVS 05: New Developments and Research Trends for Intelligent Vehicles," *IEEE Intelligent Systems*, July/Aug. 2005, pp. 10–14.
5. J. McCall and M.M. Trivedi, "Video-Based Lane Estimation and Tracking for Driver Assistance: Survey, System, and Evaluation," *IEEE Trans. Intelligent Transportation Systems*, March 2006, pp. 20–37.
6. E.C. Nelson et al., "An Embedded Architectural Framework For Interaction Between Automobiles and Consumer Devices," *Proc. IEEE Real-Time and Embedded Technology and Applications Symp.*, IEEE Press, 2004, pp. 192–199.
7. J. Waldo, "Embedded Computing and Formula One Racing," *IEEE Pervasive Computing*, vol. 4, no. 3, 2005, pp. 18–21.
8. M.M. Trivedi et al., "Occupant Posture Analysis with Stereo and Thermal Infrared Video: Algorithms and Experimental Evaluation," *IEEE Trans. Vehicular Technology*, vol. 53, no. 6, 2004, pp. 1698–1712.
9. M.M. Trivedi, T. Gandhi, and J. McCall, "Looking-In and Looking-Out of a Vehicle: Selected Investigations in Computer Vision Based Enhanced Vehicle Safety," to be published in *IEEE Trans. in Intelligent Transportation Systems*, 2006.
10. R. Bishop, *Intelligent Vehicle Technology and Trends*, Artech House, 2005.
11. S.Y. Cheng, S. Park, and M.M. Trivedi, "Multiperspective Thermal IR and Video Arrays for 3D Body Tracking and Driver Activity Analysis," *IEEE Int'l Workshop Object Tracking and Classification beyond the Visible Spectrum*, IEEE Press, 2005, pp. 132–139.
12. H. Veeraraghavan et al., "Driver Activity Monitoring through Supervised and Unsupervised Learning," *IEEE Conf. Intelligent Transportation Systems*, IEEE Press, 2005.

might avert tragedy, or it could annoy the driver to the point where he or she ignores the warning, completely eliminating the system's safety benefits. On the other hand, in some instances, an alert might not arrive when the system encounters a situation beyond its design specifications. In this case, if the driver has placed excessive confidence in the system, he or she might not be ready to respond. Information about driver intent should let the system mitigate both of these problems and improve the driver's attention allocation. However, even given a very reliable alert system, an unforeseen situation might still raise problems in the form of driver annoyance or overconfidence.

In this case, it's helpful to distinguish between driving assistance and automation. Driving assistance bypasses the dif-

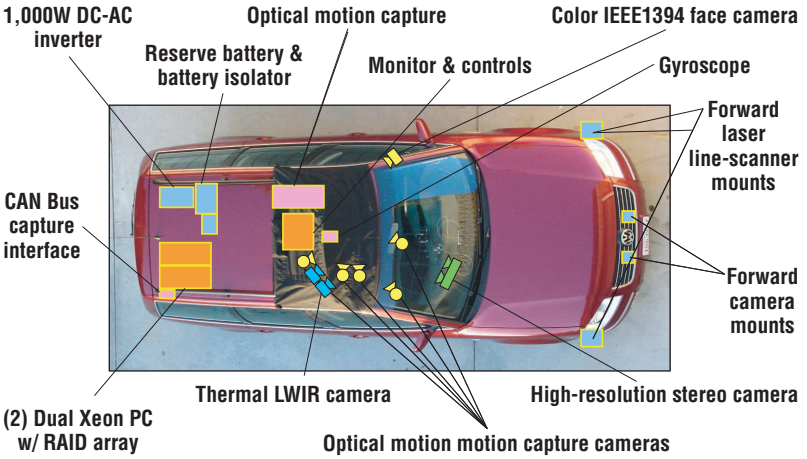


Figure 1. A pictorial diagram of the LISA-P (Laboratory for Intelligent and Safe Automobiles-Passat) testbed sensors.

ficult issue of dynamic allocation of function or the legal question of who is responsible and when. We propose keeping the driver in the loop and directly communicating the improved level of confidence that the system has in the guidance it offers. However, this requires effectively recognizing driver intent.

Driver body-pose analysis

Many motion-capture technologies already exist to recover body pose using mechanical, magnetic, and optical marker-based sensors. These sensors provide both 3 degrees-of-freedom position information of a point on the subject's body and full 6 DOF body-part position and orientation. However, such techniques require that the subject wear something during the capture, either the sensor device itself or markers that are tracked from sensors placed some distance away.

For this reason, image-based motion capture would be more applicable in a consumer car, because it doesn't require special markers or user intervention. For more than 25 years, the computer vision community has researched image-based body-pose recovery,[2] developing algorithms that rely only on passively sensed images of the driver to analyze and extract body-pose information.

However, vision-based body-pose recovery is challenging, primarily because of the wide range of configurations and appearances a human body can assume, and its tendency to occlude itself in images (self-occlusion). The problem is further complicated by the vehicular requirement for algorithms to be robust to changing illumination. We recently showed the promise of using long-wavelength infrared imaging sensors to locate exposed skin and recover hand position and pose in the presence of lighting that changes from one moment to the next.[3] This approach senses not the visible appearance but the thermal response of the hands, which hovers around a constant temperature. However, the approach must also contend with varying temperatures inside the car—a much slower source of variation in the scene—and the cost effectiveness of using multiple thermal infrared cameras.

Vision-based body-pose recovery is challenging because of the wide range of configurations and appearances a human body can assume and its tendency to occlude itself in images.

Consequently, we rely on a marker-based method of recovering poses. Such a method, if fast and accurate, can help us in two ways. First, it can provide ground-truth data for pose-recovery development, which can be invaluable in gauging a vision-based pose-recovery system's performance when comparing techniques. Second, it helps us understand the upper limits of an activity-recognition system based on pose data by providing the system with the cleanest, most complete data set. Knowing this limit helps justify dedicated development efforts toward passive and likely application-specific pose-recovery systems.

Here, we focus on the second benefit—clean data. Using driver body-pose information collected from marker-based motion capture, we developed a system to recognize and predict driver left and right intersection-turn behaviors.

Recognizing and predicting intersection turns

We characterize an intersection turn as one that requires a change in vehicle heading of approximately 90 degrees. A driver would encounter such turns when maneuvering a vehicle at T and four-way intersections as well as turnouts into side streets. These intersections have various numbers of start and end lanes, combinations of one- and two-way streets, and sometimes certain traffic controls (such as stoplights, stop signs, or yield signs).

Of course, there are many other kinds of turns—those involving freeway ramps or curves in the road, those made to avoid road obstructions, U-turns, and so forth. However, we focus on inter-

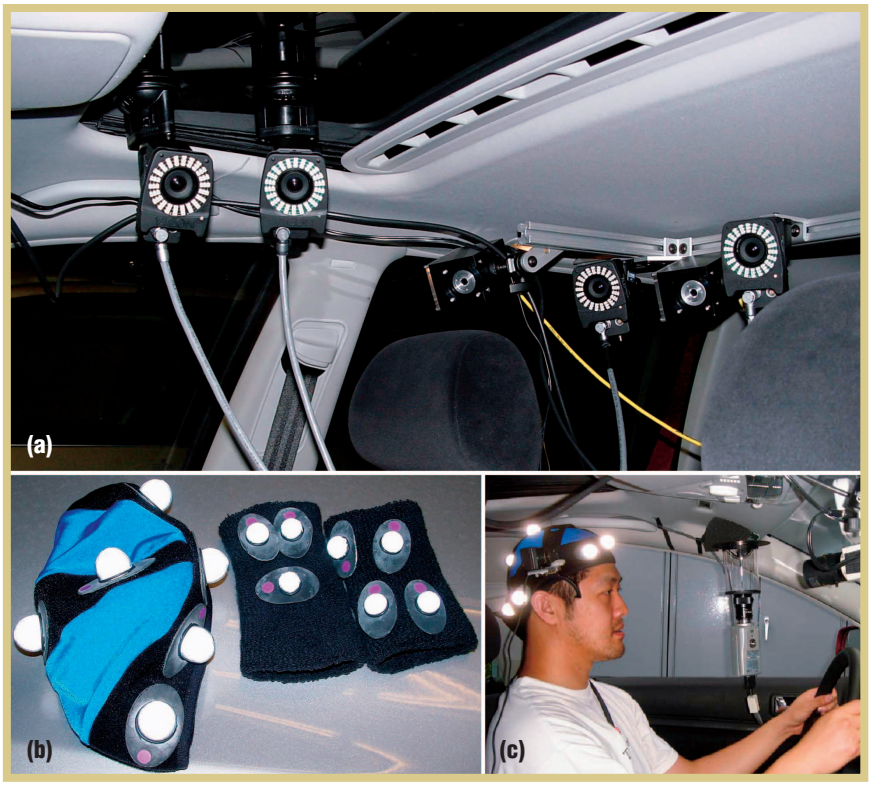


Figure 2. (a) Arrangement of the vehicle's motion-capture cameras (with near-infrared lights arranged in a ring), (b) the head and wrist markers, and (c) a driver wearing the head markers.

section turns because they're relatively prevalent and prone to involve road obstacles. There are two kinds of intersection approaches: you can stop and then start into the turn (a slow turn), or you can stop, wait at least two seconds, and then start into the turn (a halt turn). We focus on the first kind, specifically training our system to ignore all other kinds of turns and behaviors.

We collected the driver-pose data using a commercial motion-capture system that uses retroreflective markers placed on the subject and kinematics software to recover the 6-DOF head pose and 3-DOF hand positions. We installed and operated the system in the LISA-P (Laboratory for Intelligent and Safe Automobiles-Passat), a specially equipped test vehicle based on the Volkswagen Passat (see figure 1). We placed four cameras around the driver, attached to glass surfaces with a suction mount on the moon roof and windshield (see figure 2a). Then, to capture the driver's head pose, we placed six markers on his head—two along the arc from the top to the front and four around the sides (see figure 2b and 2c). To capture his left- and right-hand positions, we placed three and four markers on the left and right wrists, respectively (see figure 2b). The different numbers of markers let the system distinguish between the two hands.

At each time m, the motion-capture system collects a data vector $\mathbf{p}_m$ that comprises 12 values representing the head's position ($\mathbf{t}_{h,m}$) and orientation ($\mathbf{o}_{h,m}$) and the left and right hands' positions ($\mathbf{t}_{l,m}$, $\mathbf{t}_{r,m}$):

$$\mathbf{p}_m = \begin{pmatrix} \mathbf{t}_{h,m} \\ \mathbf{o}_{h,m} \\ \mathbf{t}_{l,m} \\ \mathbf{t}_{r,m} \end{pmatrix}$$

We model whether or not an intersection turn is or will be taking place as a linear regression of a sequence of body-pose parameters. Namely, the classifier $c(\mathbf{y}_m)$ follows the form

$$c(\mathbf{y}_m) = \mathrm{sgn}(s(\mathbf{y}_m)) = \mathrm{sgn}(\mathbf{w}^{\mathrm{T}}\mathbf{y}_m + \tau) \quad (1)$$

where the coefficient $\mathbf{w}$ is found using sparse Bayesian learning,[4–6] the threshold τ is used to tune to a desired false-alarm and missed-detection rate, and sgn is the sign function. The feature vector $\mathbf{y}_m$ is a collection or window of pose measurements of the last M samples up to the current sample m:

$$\mathbf{y}_m = \begin{pmatrix} \mathbf{p}_m \\ \mathbf{p}_{m-1} \\ \vdots \\ \mathbf{p}_{m-M+2} \\ \mathbf{p}_{m-M+1} \end{pmatrix}$$

We trained the regression function $s(\mathbf{y}_m)$ to produce a value of +1 if the input pose sequence is the result of the left (or right) intersection turn, and –1 if not. This is achieved by learning the weights $\mathbf{w}$ that would produce this behavior from positive and negative examples of intersection turns using sparse Bayesian learning.[4] Many other possible approaches exist for solving this classification problem, including hidden Markov models,[7] support vector machines,[8] or adaptive boosting with weak classifiers. However, the attractive aspect of sparse Bayesian learning is the resulting sparcity of the solutions of $\mathbf{w}$. In other words, most elements of $\mathbf{w}$ will be zero or close to zero, the effect of which is that the classifier ignores certain pose measurements at certain times in deciding whether the sequence observed indicates an impending turn. This saves computation cycles, and when elements of $\mathbf{w}$ associated with particular measurements are zero for all times, we can even avoid collecting those measurements entirely. (The classifiers we describe here generated coefficients, so we couldn't ignore entire measurements. However, when the classifier is trained with other data, such as other car dynamic measurements, this attribute of sparse Bayesian learning

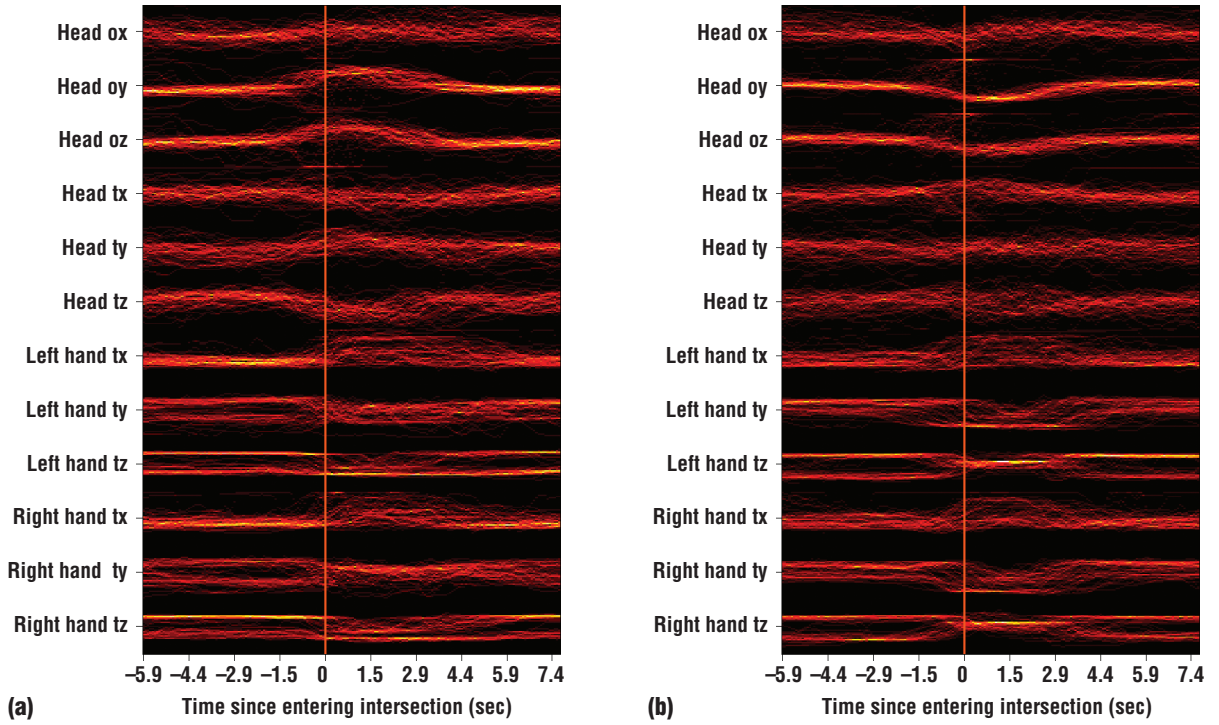


Figure 3. A histogram of body-pose time-sequences, with time aligned at the start of the intersection turn (marked by the vertical line). The sequences shown are $(\theta_x, \theta_y, \theta_z)$ orientation values of the head and (x, y, z) position coordinates of the head, left hand, and right hand for (a) left turns and (b) right turns. (Values fluctuate in a perceivable pattern during the intersection turn itself and about 1 second before it. The origin of the coordinate system is located at one of the four motion-capture cameras.)

TABLE 1
Collected intersection-turn data.

	Left turn	Right turn
Maximum (sec)	9.95	7.30
Average (sec)	5.89	4.06
Minimum (sec)	2.14	2.21
Standard deviation (sec)	1.59	1.24
Total number of turns	68	80

lets us gauge the relevance of adding those additional measurements.[6])

We can vary the threshold τ to obtain the desired detection and false-alarm rate. This parameter is significant because it influences the user's level of trust in the system. Suppose the system generates driver alerts when it detects that a maneuver at that moment is dangerous. A high false-alarm rate implies that the system often unnecessarily alerts the driver of danger. The positive side to this is that a high false-alarm rate usually accompanies a high correct-detection rate, implying that whenever a dangerous maneuver is about to be made, the system will detect it. On the other hand, a low false-alarm rate and high correct-detection rate implies that whenever the system predicts a dangerous maneuver, it will likely be true, but it will miss many other dangerous maneuvers. Both extremes will diminish the system's effectiveness by annoying the driver with unnecessary warnings or by missing several dangerous maneuvers. This corresponds to the issue of proactivity versus transparency in pervasive computing. We must find a balance by using threshold τ to tune the system.

Positive samples of $\mathbf{y}_m$ for training the classifier are extracted from motion data where m is aligned with the decision time. The decision time D is when we want the classifier, through training, to determine whether or not the driver will be or is making a turn. We speak of the decision time in relative terms from the start of an intersection turn, when the vehicle has just entered the intersection. We can also interpret the inverse (or negative) of the decision time as the desired

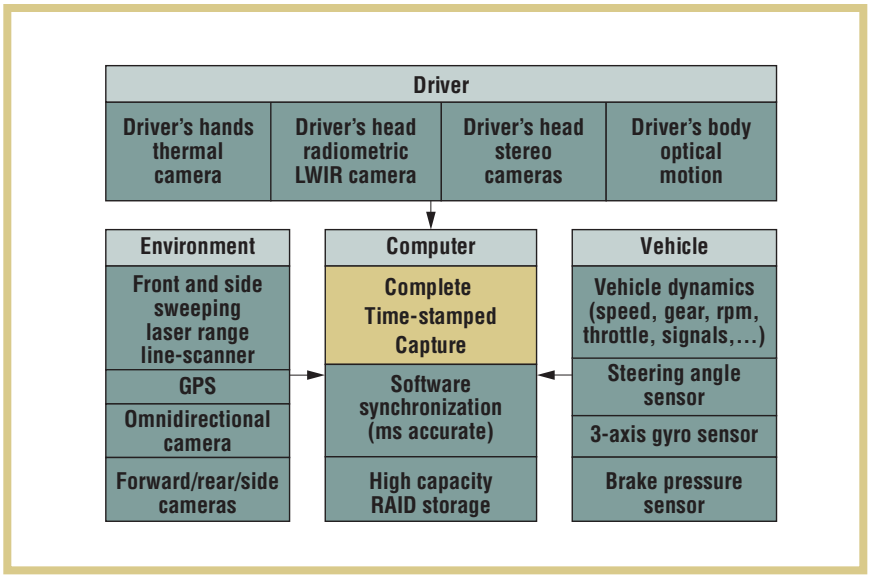


Figure 4. A block diagram of the LISA-P testbed sensors.

amount of lead time, but here we use decision time.

It makes sense to use all the available preparatory motion data up to the time when a decision is desired. So, the desired decision time also determines the length of the window M. We fix the window length to 3 seconds for negative decision times ($D < 0$) and $3 + D$ seconds for positive decision times ($D > 0$). Negative samples of $\mathbf{y}_m$ for training the classifier are of the same length M. They're extracted everywhere else in the motion sequence spaced out at regular intervals, which we set at every three samples, or 0.2 seconds. We can train a different classifier for different decision times and window lengths.

We told our subject to drive the LISA-P along a route on one- to three-lane local city streets containing slow-turn intersections. We collected pose data for 148 slow intersection turns over 3.5 hours—68 left and 80 right slow turns. The average slow turn duration was 5.89 and 4.06 seconds for left and right turns, respectively. (See table 1 for the statistics of all the turns.) Figure 3 shows a 2D histogram of pose sequences aligned at the start of each intersection turn. You can see a clear pattern of movement at each turn as well as a pattern distinguishing between left and right turns. The body-pose measurements made for left and right turns appear in the plots to swing in opposite directions.

Experimental testbed and results

The LISA-P testbed is designed to obtain complete coverage of the vehicle's surroundings, interior, and state for extended time periods from various modular sensing systems. Sensors include the motion-capture system previously discussed and a suite of vision and other multimodal sensors, including vehicle state information such as steering angle, throttle, brake pressure, and so on. Figure 4 shows an overview of the LISA-P's sensory capabilities. Here, we first describe classifier performance using driver-pose information and then compare it to steering-angle data.

To measure the performance of our pose-based intersection-turn intent recognition and prediction system, we generated a series of receiver-operator-characteristics (ROC) curves for the slow-left and slow-right turn classifiers while varying the decision time D. We used half of the positive and negative samples for training, and the other half for testing. The series of ROC curves shown in figure 5 present a complete picture of the classifiers' performance on the testing data for classifiers trained for a specific decision time. A good classifier can operate at a high true-positive rate and a low false-alarm rate.

In both the left- and right-turn cases, the classifiers trained to make a decision at $D = 1.5$ seconds (meaning they recognized turn intent about one-third of the way into the average turn) produced the highest true-positive rate. Their rates were 70 and 85 percent for the left and right turns, respectively, with a 10 percent false-alarm rate. The true-positive rate increases to 82 and 96 percent with a 20 percent false-alarm rate. Operating in a predictive mode with decision time $D = 0$ seconds (that is, just before the vehicle enters the intersection), the classifier using only pose data correctly recognized positive samples 82 and 75 percent of the time for left and right turns and produced false alarms from the negative samples 20 percent of the time.

A good classifier can operate at a high true-positive rate and a low false-alarm rate.

Figure 6 shows the true-positive and true-negative rates of classifiers trained at several decision times. Each curve in the figure represents a particular operating point along the ROC curve. These four operation points are when the false-alarm rate equals the missed-detection rate (1 minus the true positive rate) and when the false-alarm and missed-detection rates each equal 10,

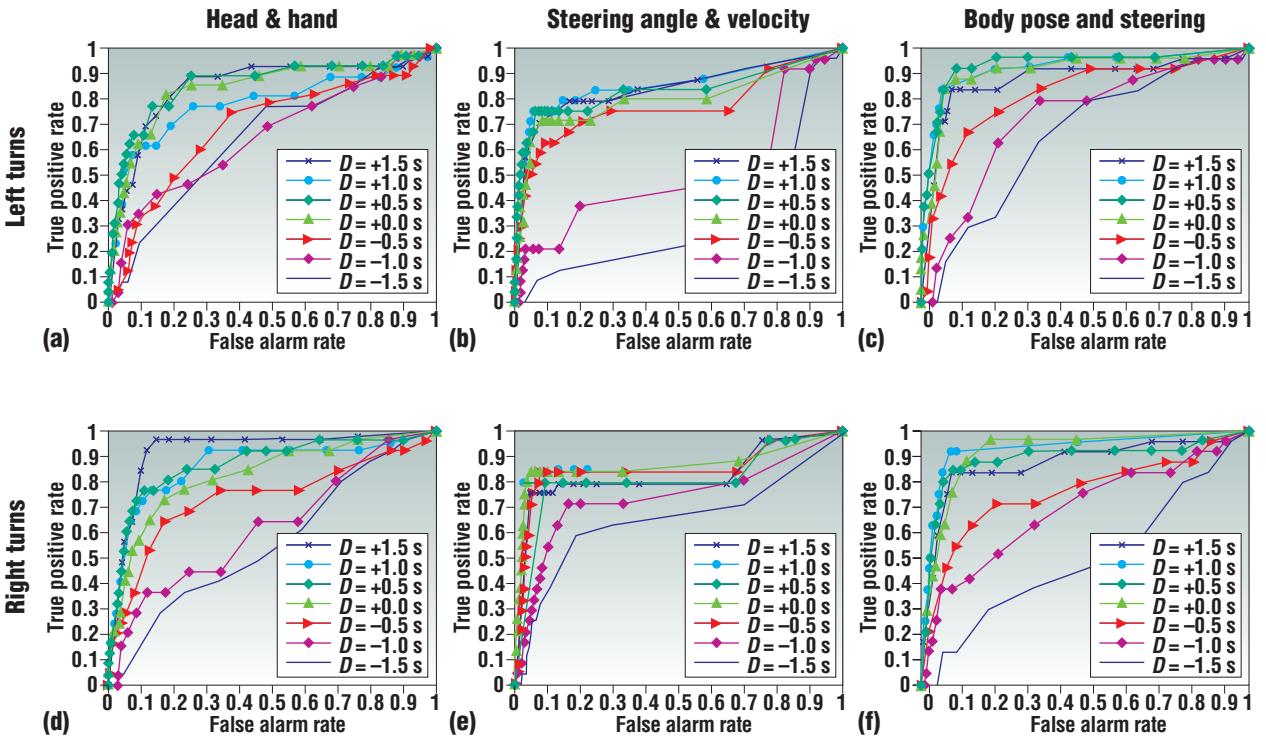


Figure 5. Receiver-operator-characteristics curves of (a, b, and c) left-turn and (d, e, and f) right-turn classifiers using (a, d) both head and hand data, (b, e) steering angle and velocity data, and (c, f) body-pose and steering data together. Each ROC curve is the performance of a classifier that is trained to make a decision at –1.5, –1, –0.5, 0, 0.5, 1, and 1.5 seconds from the start of the intersection turn. (*D* stands for decision time.)

15, and 20 percent. The plot with equal false-alarm and missed-detection rates is the same in both true-positive and true-negative plots. It acts as a baseline to relate the two plots for a given classifier.

Figure 6 clearly shows the classifier's predictive power at several decision times before and after the start of the turn, with the best-performing decision times just after the start of the turn. The classifier's true-positive and true-negative rates at various decision times are the results of training the classifier with positive and negative samples aligned with the same decision times. The true-positive rate plots show how well the classifier correctly classifies driver turn intent at various decision times from –3 to 2 seconds from the start of the turn. Highest performing periods are after the start of the turn. The true-negative rate plots show how well the classifier correctly classifies negative instances when there is no turn intent with the same range of decision times.

The classifier functions like a filter to the pose-measurement time sequence. Figure 7 shows the statistics of the regression response of equation 1 at each point in time surrounding the start of the turn. The figure shows the sigmoid of the regression function $s(\mathrm{y}_m)$ for classifiers trained at 0 and –0.5 seconds decision times for the left- and right-turn intents, respectively. The decision boundaries are marked by the vertical dotted line at 0 and –0.5 seconds, times relative to the start of the turn. The statistics of the regression response for the several seconds surrounding the start of the turn reveal qualitatively the distribution of regression response over all the test samples. They are denoted by the inner and outer shaded regions, which represent the standard deviation and minimum-maximum values of the response at that time, both of which are centered at the dotted line representing the regressor's mean response. The threshold τ is the value that would yield equal false-alarm and missed-detection rates.

To illustrate how informative head pose and hand positions can be, we trained classifiers using head pose alone and hand positions alone. Figure 8 shows that hand position provides the best classifiers from –1 to 2 seconds deci-

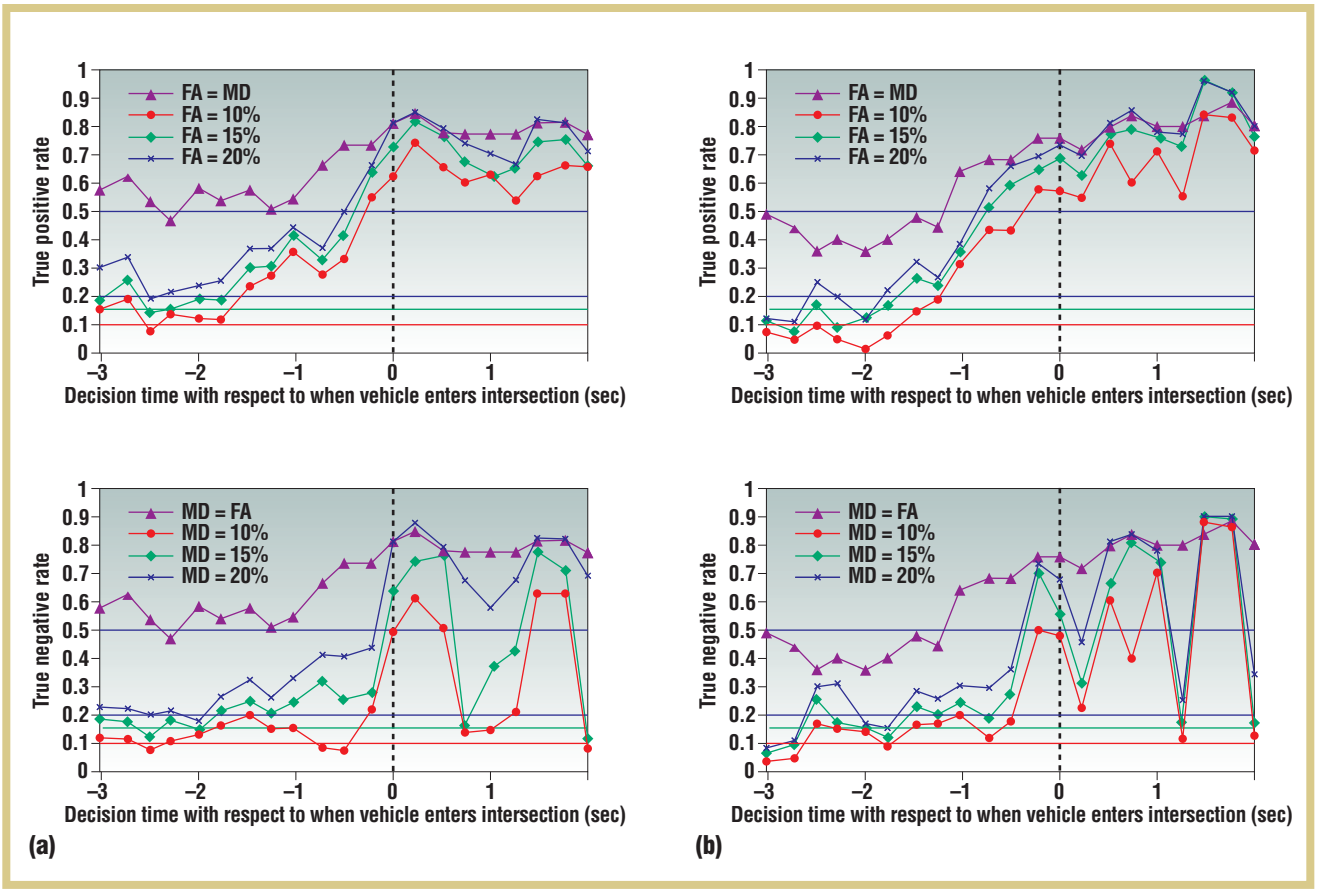


Figure 6. True-positive (top) and true-negative (bottom) rates vs. decision-time plots illustrate the predictive power of the available data for (a) slow-left and (b) slow-right turns over various decision times. For each decision time, we generated a different set of positive and negative samples from the 148 skew turns to train the classifier. Horizontal lines at 0.1, 0.15, 0.2, and 0.5 true-positive or true-negative rates indicate the levels that can be achieved by random guessing for a false alarm (FA) rate = 10 percent, 15 percent, 20 percent, and the missed-detection (MD), respectively, and for the MD rate = 10 percent, 15 percent, 20 percent, and FA, respectively.

sion times, while the head pose provides best classifiers between –3 and –1 seconds decision times. Together, they produce the best classifiers, but with hands contributing to the best rates with –1 second or later decision times.

We also wanted to illustrate that there's still something to be gained from looking beyond sensor data that's available in cars today, such as steering data, which consists of the steering angle as well as angular-velocity data. So, we compared the performance of a turn-intent classifier trained on only pose data with a classifier trained on only steering angle data captured through the vehicle's controller-area-network bus. We also examined the performance of a classifier using both driver-pose and steering-angle data.

Figure 5 shows the ROCs of pose-only, steering-only and pose-plus-steering classifiers. Figure 9 shows the true-positive rate of the pose-only, steering-only, and pose-plus-steering classifiers trained at various decision times. These curves represent operating points where the false-alarm rate equals the missed-detection rate.

The performance of both the steering and pose-only data are comparable except between –3 and –1 seconds decision times, during which time the steering-only classifier consistently outperformed pose-only at a better than 80 percent true-positive rate (corresponding to less than 20 percent missed-detection rate). The classifier trained with both pose and steering data produced the best true-positive rates between –1 and 1.5 seconds decision times at better than 85 percent. At some decision times, this classifier achieved a full 10 percent better performance than the other two classifiers. Between 0 and 1.25 seconds, the classifier achieved better than 90 percent true-detection rate with a less than 10 percent false-alarm rate. Clearly, adding pose data increased accuracy in driver intent predication at several decision times.

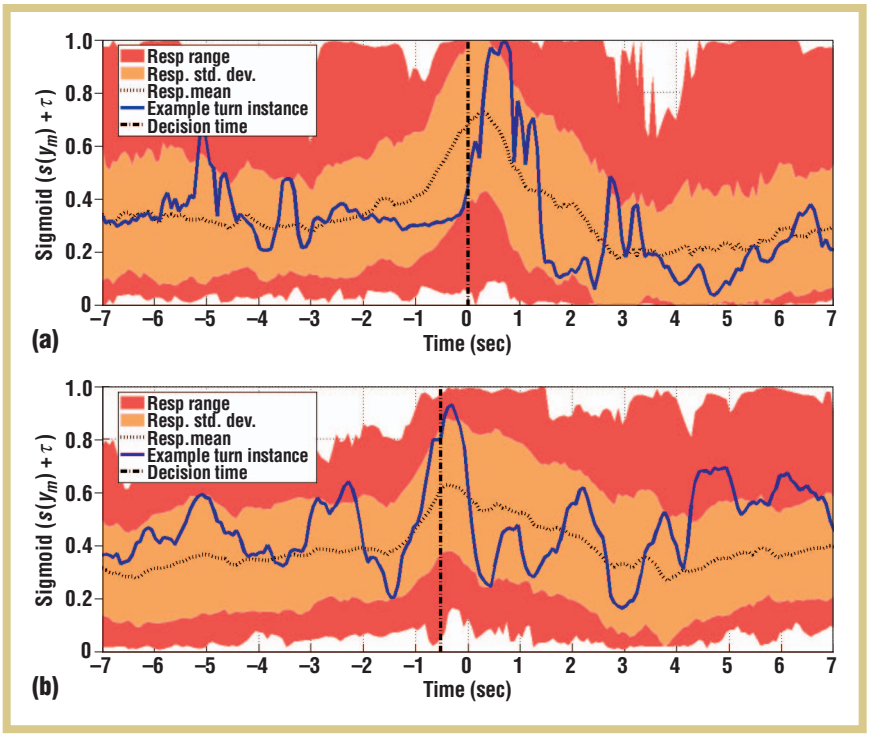


Figure 7. Responses over time of a regressor trained for –0.5 s and 0 s decision times (just as the vehicle enters the intersection and 0.5 seconds before) using driver head and hand pose only for (a) slow-left and (b) slow-right turns. We tuned the chosen threshold to produce equal false-alarm and missed-detection rates. The inner and outer shaded regions represent the standard deviation of regression responses from the mean and range of values that the regression response takes on at each point in time.

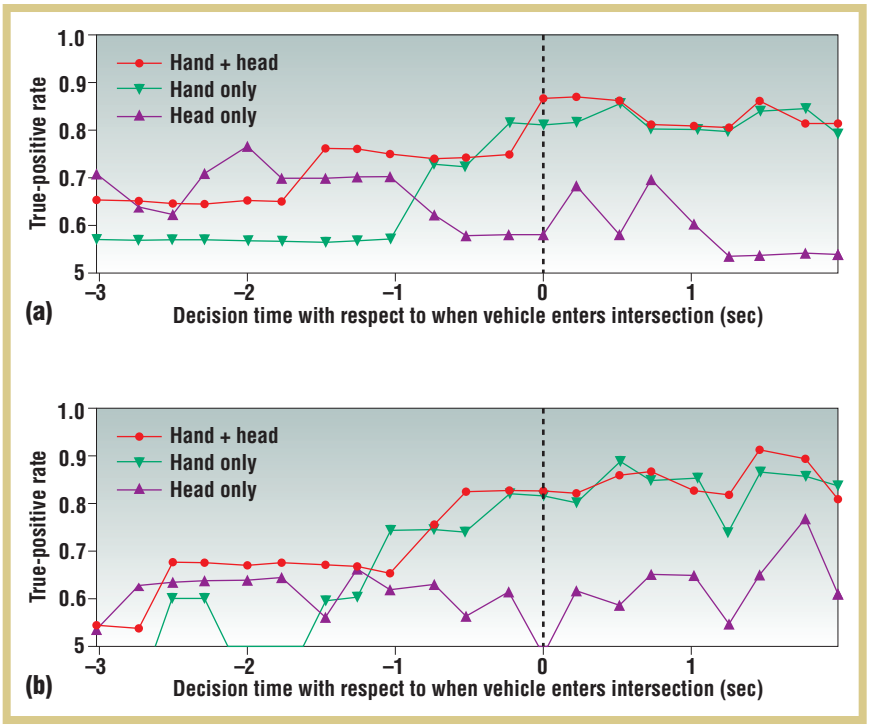


Figure 8. True-positive rates vs. decision time plots illustrate the predictive power of classifiers using hand and head data, hands-only data, and head-only data for (a) left and (b) right turns. Head and hands data complement each other (head data for before the turn, hands data during the turn) to create a classifier better than either alone. For each curve, each point represents the operating point where FA = MD. This implies that a 0.5 true-positive rate is a level of performance that can be achieved by random guessing.

The environment, vehicle, and driver comprises the three main components of the driver-vehicle system. Overall awareness of these three factors in the driving scenario allows for intelligent intervention in mitigating a dangerous situation. An intelligent vehicle aware of its dynamic surroundings and the driver's behavior would allow for a timely alert or even corrective action to avert a collision. More generally, intelligent driver support systems provide an interesting application domain for addressing some of the challenging multidisciplinary research problems in pervasive computing,[1] including that of multimodal interactions, human-centric hardware design, and algorithms for proactive behavior. P

ACKNOWLEDGMENTS

We thank Joel McCall for his invaluable advice and assistance with sparse Bayesian learning and technical aspects of the data capture equipment. A grant from the University of California Discovery's Digital Media Innovation Program and the Electronics Research Laboratory of Volkswagen of America made this research possible. We also appreciate the insightful and constructive comments made by the reviewers.

REFERENCES

1. M. Satyanarayanan, "Pervasive Computing: Vision and Challenges," *IEEE Personal Comm.*, vol, 8, no. 4, 2001, pp. 10–17.

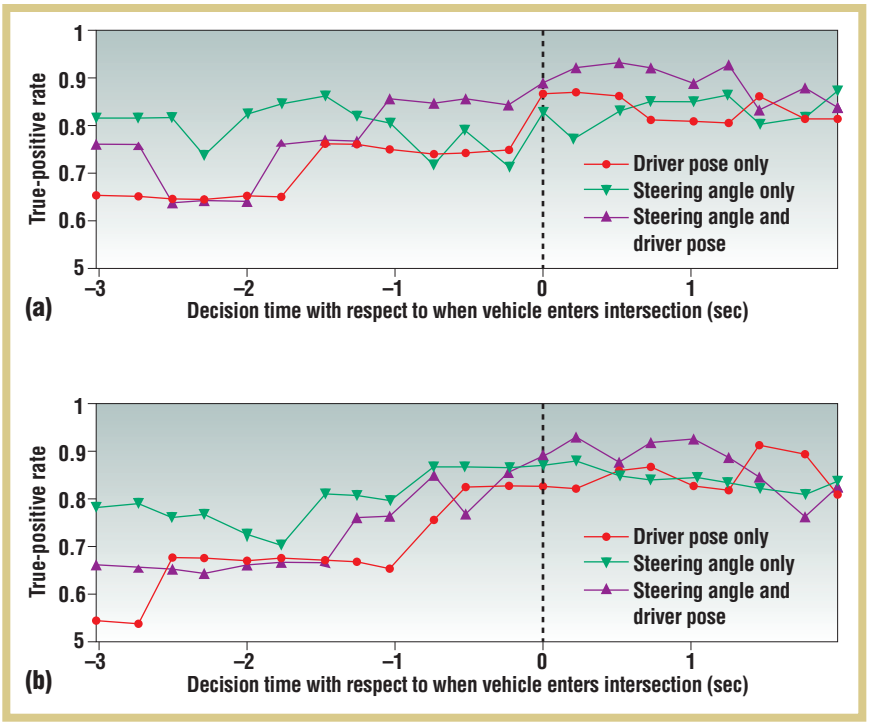


Figure 9. True-positive rates vs. decision-time plots illustrate the predictive power of classifiers using driver-pose-only (head and hands), steering angle only, and the two in combination for (a) left and (b) right turns. For each curve, each point represents the operating point, where FA = MD.

2. T.B. Moeslund and E. Granum, "A Survey of Computer Vision-Based Human Motion Capture," *Computer Vision and Image Understanding*, vol. 81, no. 3, 2001, pp. 231–268.

3. S.Y. Cheng, S. Park, and M.M. Trivedi, "Multiperspective Thermal IR and Video Arrays for 3D Body Tracking and Driver Activity Analysis," *IEEE Int'l Workshop Object Tracking and Classification beyond the Visible Spectrum*, IEEE Press, 2005, pp. 132–139.

4. M.E. Tipping, "Sparse Bayesian Learning and The Relevance Vector Machine," *J. Machine Learning Research*, May 2001, pp. 211–244.

5. J. McCall and M.M. Trivedi, "Human Behavior Based Predictive Brake Assistance," *Proc. IEEE Intelligent Vehicles Symp.*, IEEE Press, 2006, pp. 8–12.

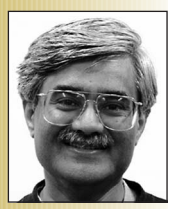
6. J. McCall et al., "Lane Change Intent Analysis Using Robust Operators and Sparse Bayesian Learning," to be published in *IEEE Trans. Intelligent Transportation Systems*, 2006.

7. N. Kuge, T. Yamamura, and O. Shimoyama, "A Driver Behavior Recognition Method Based on a Driver Model Framework," *Soc. Automotive Engineers Publication*, vol. 109, no. 6, 2000, pp. 469–476.

8. H.M. Mandalia and D.D. Salvucci, "Using Support Vector Machines for Lane-Change Detection," *Proc. Human Factors and Ergonomics Soc. Conf.*, 2005.

the AUTHORS

Shinko Yuanhsien Cheng is a researcher at the Computer Vision and Robotics Research Laboratory and a PhD candidate in computer vision at the University of California, San Diego. His research interests include computer vision, human-machine interface, articulated body modeling, volumetric scene reconstruction, human body-pose and structure recovery, human activity recognition, intelligent transportation systems, and intelligent vehicles. He received his MS in electrical engineering from the University of California, San Diego. Contact him at the Computer Vision and Robotics Research Lab., Univ. of California, San Diego, Science and Eng. Research Facility Rm. 150, 9500 Gilman Dr., MC 0434, La Jolla, CA 92037-0434; sycheng@ucsd.edu.



Mohan Manubhai Trivedi is a professor of electrical and computer engineering, the founding director of the Computer Vision and Robotics Research Laboratory, and leader of the Intelligent Transportation and Telematics Layer of the California Institute for Telecommunication and Information Technology (Cal-IT2) at the University of California, San Diego. His research interests include intelligent systems, computer vision, intelligent (smart) environments, intelligent vehicles and transportation systems, and human-machine interfaces areas. He has established the Laboratory for Intelligent and Safe Automobiles (LISA) at UCSD to pursue a multidisciplinary research agenda. Contact him at the Computer Vision and Robotics Research Lab; Univ. of California, San Diego, Science and Eng. Research Facility Rm. 150, 9500 Gilman Dr., MC 0434, La Jolla, CA 92037-0434; mtrivedi@ucsd.edu.