Reasoning-Based Framework for Driving Safety Monitoring Using Driving Event Recognition

Bing-Fei Wu, Fellow, IEEE, Ying-Han Chen, Chung-Hsuan Yeh, and Yen-Feng Li

Abstract—With the growing concern for driving safety, many driving-assistance systems have been developed. In this paper, we develop a reasoning-based framework for the monitoring of driving safety. The main objective is to present drivers with an intuitively understood green/yellow/red indicator of their danger level. Because the danger level may change owing to the interaction of the host vehicle and the environment, the proposed framework involves two stages of danger-level alerts. The first stage collects lane bias, the distance to the front car, longitudinal and lateral accelerations, and speed data from sensors installed in a real vehicle. All data were recorded in a normal driving environment for the training of hidden Markov models of driving events, including normal driving, acceleration, deceleration, changing to the left or right lanes, zigzag driving, and approaching the car in front. In addition to recognizing these driving events, the degree of each event is estimated according to its character. In the second stage, the danger-level indicator, which warns the driver of a dangerous situation, is inferred by fuzzy logic rules that address the recognized driving events and their degrees. A hierarchical decision strategy is also designed to reduce the number of rules that are triggered. The proposed framework was successfully implemented on a TI DM3730-based embedded platform and was fully evaluated in a real road environment. The experimental results achieved a detection ratio of 99% for event recognition, compared with that achieved by four conventional methods.

Index Terms—Driving events, driving safety, fuzzy logic, hidden Markov models (HMMs).

I. INTRODUCTION

STUDY by the National Highway Traffic Safety Administration reported that each year, approximately 56 000 car accidents are caused by driver fatigue, in which approximately 1500 drivers die. In Taiwan, a traffic accident occurs every 2 min on average, whereas drunk driving and other human errors were the prime causes of accidents in the decade studied [1]. With advances in automotive electronics in recent years, a wide range of automotive safety systems have been released, such as driver blind-spot detection and a 360° panorama imaging system [2], [3]. These driving assistance systems focus on the detection of certain events, such as lane departure, or the approach of vehicles in the front/back/blind spots. They

B.-F. Wu, Y.-H. Chen, and C.-H. Yeh are with the Department of Electrical Engineering, National Chiao Tung University, Hsinchu 300, Taiwan (e-mail: winand@cssp.cn.nctu.edu.tw).

Y.-F. Li is with CSSP Inc., Hsinchu 300, Taiwan.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TITS.2013.2257759

provide drivers with warnings when such events are detected [4], because a timely warning may prevent a possible accident. The occurrence of an accident is related to many factors. We believe that if dangerous events frequently occur, there is a high probability that an accident will occur. For drivers who are responsible for the transportation of people or goods, an accident has a more serious impact than it does for individuals. If we could choose drivers with better driving performance, i.e., no frequent occurrence of dangerous events, this would facilitate the safety and protection of both people and goods. For this reason, there has been an increase in research into the monitoring of a driver's performance to prevent potential risks.

A driver's performance can be evaluated using several indexes. People intuitively understand poor driving behavior when they feel that a driver is performing dangerous maneuvers that make them uncomfortable. However, it is difficult to define what kind of behavior is dangerous or the danger level that it represents. Because exact danger patterns have not been established, the danger level of a given behavior may change when it takes place in different environments. Although the time at which a detected event occurs can be recorded, it is still difficult to provide a reasonable value of the danger level to evaluate the driver's performance in terms of any single event or when it occurs. Recent years have seen a growing body of research on driving safety, and the methods used to determine the degrees of dangerous driving are broadly divided into three categories, according to the subjects being studied: 1) the states of the driver, 2) the states of the vehicle, and 3) the hybrid [5], [6].

A. Driver's States

The first category considers the danger for which the driver is responsible. Different danger levels are determined by the driver's state using various physiological signals and their implications. In such cases, driver fatigue, distraction, and drunkenness are often a cause of danger. If a system could identify these states, system warnings might be able to awaken a driver in time to correct his/her course, and possible accidents might be avoided. This research approach thus focuses on how to effectively detect the driver's state.

Some reports in the literature employ physiological signals, such as the heart rate [7], [8], electroencephalogram (EEG) [9]–[12], electrocardiograph (ECG) [13], and the respiration rate [14], to determine the driver's state. In [7], a new calculation method for respiratory sinus arrhythmia and Mayer wave-related sinus arrhythmia is derived from an analysis of heart rate variability, to evaluate the driver's mental stress and drowsiness.

Manuscript received September 6, 2012; revised February 8, 2013; accepted April 2, 2013. Date of publication April 30, 2013; date of current version August 28, 2013. This work was supported by the National Science Council (NSC) under Grant NSC 101-2221-E-009-099. The Associate Editor for this paper was H. Dia.

Jap *et al.* [9] assessed four EEG activities, namely, delta, theta, alpha, and beta, during a monotonous driving session. The results have implications for the detection of fatigue, based on an increase in the ratio of slow- to fast-wave EEG activities over time. Chua *et al.* [13] combined ECG and photoplethysmogram measurements to estimate psychomotor vigilance by establishing multiple linear regression models. Yang *et al.* [15] proposed a driver fatigue recognition model based on a dynamic Bayesian network that uses physiological features (ECG and EEG), and they applied the hidden Markov model (HMM) to compute the dynamics of the network at different points in time. To acquire the given signals, however, one or more intrusive sensors must be attached to the driver's body. This requirement makes most drivers unwilling to use the system and is impractical in most situations.

To avoid these shortcomings of intrusive sensors, the other approach to research uses cameras to identify driver fatigue and distraction via the recognition of characteristic facial expressions. Single or multiple facial features, such as the percentage of eye closure (PERCLOS), eye closure duration, blink frequency, nodding frequency, face position, and fixed gaze, are used to characterize a driver's state [16]-[19]. Albu et al. [20] proposed an approach that focuses on a single visual cue and uses a custom-designed template-matching algorithm for online eye-state detection of fatigued drivers. Qiang et al. [21] described a driver fatigue monitor that recognizes eyelid movement, gaze movement, head movement, and facial expression. A probabilistic model was then developed to model human fatigue based on visual cues. Ueno et al. [22] developed a system for drowsiness detection that recognizes whether a driver's eyes are open or closed, using image processing technology. Their preliminary evaluations gave promising results, which showed that the system's performance is comparable to that of techniques that use physiological signals. However, the variations in light and shadow during the day affect the visual appearance of the driver, which makes it a difficult challenge to use computer vision techniques to reliably obtain accurate and robust results.

Another approach involves finding the correlations between the car performance and the driver's state [23]–[25]. Eskandarian and Mortazavi [26] used an artificial neural network (ANN)-based algorithm to detect drowsiness using only input from the steering wheel. However, it is not easy to acquire such information from a vehicle without the use of a controller area network (CAN) bus or additional supports.

B. Vehicle's States

The second approach deals with the behavior of vehicles and not of drivers. The features of interest in this approach include the parameters of the host vehicle and of the environment, such as lateral positions, accelerations/decelerations, the distance to the vehicle in front, and the distance to the lane markings [27], [28]. These features also indirectly indicate the driver's state and can be extracted by various sensors. Many researchers have conducted studies along these lines to develop driving safety systems. Liu *et al.* [29] proposed a method of hazardousevent detection that employs object tracking and a finite-state machine. The tracking results are mapped to a region called a driving environment state map and are then used to predict the behavior of the tracking vehicle. Inata *et al.* [30] proposed the modeling of longitudinal driving behavior based on urban driving data of human pedal operation for longitudinal vehicle control. Huang [31] constructed a long- and short-term model for filtering the variations of parameters, and the results are classified to define different types of dangerous driving situations. However, there are several challenges in developing a safety monitoring system based on the identification of vehicle behaviors. For rule-based approaches, it is difficult to design a comprehensive set of rules that cover all dangerous behaviors.

For this reason, some researchers have used statistical methods to determine the dangerous behaviors using pattern recognition techniques. Zhou et al. [32] proposed a discriminative learning approach, i.e., conditional random field, which combines multichannel sequential data to detect unsafe driving patterns, using semisupervised learning algorithms. Ning et al. [33] proposed an approach to learning the danger-level function. The danger level is treated as an expected future reward, and temporal difference learning is used to learn the function to approximate the expected future reward. Wang et al. [34] introduced a dangerous-driving warning system that uses statistical modeling to predict driving risks. In this system, a semisupervised learning method utilizes both labeled and unlabeled data to build an appropriate danger-level function. Aoude et al. [35] developed algorithms based on support vector machines (SVMs) and HMMs for estimating driver behavior at road intersections, which are combined with Bayesian filtering to classify drivers as being compliant or in violation. Patterns of dangerous behavior are difficult to define, and the boundaries of the patterns may not be so clear. Another research group has used regression techniques to identify levels of danger [36]. In this case, however, it is also difficult to explain the influence of each parameter in the regression model.

C. Hybrid

Hybrid means that states of both the driver and the vehicle are simultaneously considered. Sathyanarayana *et al.* [37] tried to detect distraction by using data from the CAN and motion sensors (accelerometer and gyroscope). The collected data were analyzed using principal component analysis and linear discriminant analysis to reduce the number of features. Then, a k-nearest neighbor classifier was trained and verified. The results showed that this system could yield an accuracy of over 90% for distraction detection. However, not every related work in this category showed the comparison detection results of using the single state and the hybrid. It is difficult to conclude that using the approach of the single state is definitely worse than the approach of the hybrid.

In this paper, we propose a reasoning-based framework to provide drivers with an intuitively understandable indicator that alerts them to a definite level of danger, which is based on the recognition of different driving events. There are two stages in our framework. The first stage is the recognition of the following seven driving events, using HMMs: normal driving (ND), acceleration (ACC), deceleration (DEC), changing left (CL), changing right (CR), zigzag driving (ZD), and approaching the car in front (AFC). In the second stage, a fuzzy inference system (FIS) is introduced that produces a danger-level indicator that alerts the driver to dangerous situations. The higher the value indicated, the worse the status of the driver's situation. The data that are used come from a camera, an accelerometer, and a global positioning system (GPS) receiver, which provide information regarding lane bias, the distance to the car in front, frontal and lateral accelerations, and the velocity of the host vehicle. We sought to treat the process in a clearly analytical manner. The first stage is a classification-based problem, which requires a well-designed approach to obtain a good detection ratio. The second stage relies on an expert system that we expect can offer drivers a clear and reasonable explanation of the meaning of the danger level that is indicated.

The main contribution of the proposed system is that it provides drivers not only with a simple indicator of the danger level but also with information that defines the present driving event. This system is, thus, particularly useful for management applications, as it can use easily installed sensors that have interfaces that are nonintrusive for either people or vehicles, fostering increased willingness to apply the system. Moreover, all training and testing data were acquired from a real vehicle in a real environment, and the framework was implemented on an embedded platform to evaluate its performance.

The rest of this paper is organized as follows: Section II introduces the design of the vehicle behavior model and the usage of selected sensors. The details of driving event recognition and the logic that governs the different danger levels are addressed in Sections III and IV, respectively. The experimental results and comparisons of performance are presented in Section VI. Finally, our conclusions are presented in Section VI.

II. SETUP OF THE VEHICLE BEHAVIOR MODEL

The act of driving consists of many complicated behaviors and factors. To easily and constructively describe and analyze driving behavior, a vehicle behavior model must first be developed. In surveying past research, we found that most driving situations can be described in terms of three aspects, namely, longitudinal, lateral, and car-following behaviors. As shown in Fig. 1, longitudinal behaviors characterize the longitudinal variation of vehicles, such as acceleration and deceleration, whereas lateral behaviors explain the lateral variation of vehicles. In general, there is not just a single car on the road, and most accidents are caused by collisions between vehicles. For this reason, car-following behaviors are introduced to account for the variable distance between one car and the car in front of it. We believe that the combination of these three aspects can be used to model most driving behaviors and provide sufficient information to infer the level of danger. ACC, DEC, CL, CR, ZD, and AFC were selected as representative behaviors that cover the three aspects. The details of these behaviors are listed in Table I. In addition, the given behaviors are also referred to as "driving events" in this work, and the recognition of these driving events is the most important task in the first stage of the proposed framework.



Fig. 1. Vehicle behavior model.

TABLE I Driving Behavior Categories

Longitudinal Lateral behaviors		Car-following behaviors
behaviors		
Normal driving	Normal driving	Normal driving
Acceleration	Changing left	Approaching the car in front
Deceleration	Changing right	
	Zigzag driving	

Unlike most research, which uses a simulator to obtain the data needed for analysis and evaluation, the proposed framework is designed to be a practical system. Accordingly, the selected sensors needed to be easily acquired and installed. For this reason, we chose a camera, an accelerometer, and a GPS receiver to collect the needed data. Most people are familiar with these different kinds of sensors, which are even available on their mobile phones. The driving events in the longitudinal and lateral behaviors are relative to the velocity and the acceleration of the host vehicle, which can be obtained from the accelerometer and the GPS receiver. The accelerometer is set on the dashboard, and the x-axis and the y-axis of the sensor coordinate system are aligned with the longitudinal and lateral directions of the vehicle, respectively. The events of lane changing and approaching the car in front, however, are related not only to the host vehicle but to the environment as well, i.e., the position relative to the lane markings and the vehicle in front.

Detection of lane markings and vehicles has been a field of interest in the development of advanced driver assistance systems. Many vision-based solutions have reported robust performance, and in many applications, they have been found to outperform other techniques [4], [28]. Therefore, to efficiently extract the needed parameters, this study has employed the results of lane markings and vehicle detection that appeared in our previous research [38], [39]. After correctly detecting the lane markings and the vehicle in front using image processing technology, the system can acquire values for the lane bias and the distance to the car in front, which are the major features related to lane changing and approaching the car in front, as shown in Fig. 2. The lane bias is derived from the lane markings, whose meaning can be clarified, as shown in Fig. 2(a). The solid line represents the center line in the complete image, and the dashed line represents the center line within the lane. Thus, the lane bias is the difference between the solid line and the dashed line.

III. DRIVING EVENT RECOGNITION

The major task in the first stage is to facilitate driving event recognition by HMMs, including the training and the



Fig. 2. Detection results of lane markings and the vehicle in front. (a) Significant lane bias. (b) Short distance to the car in front.



Fig. 3. Flowchart of the recognition procedure.

evaluation. A flowchart of the recognition procedure is shown in Fig. 3. The data collected from sensors include lane bias, the distance to the car in front, frontal/lateral accelerations, and the velocity of the host vehicle. The lane bias and the distance to the car in front are acquired by processing captured camera images that are taken at 30 frames/s. The three-axis acceleration value is provided by the accelerometer at a sampling rate of 50 Hz, while the velocity is transmitted by the GPS receiver every second. Because the data sources have different input rates, they have to be preprocessed before executing the next step.

A. Data Preprocessing

Data preprocessing consists of four subprocesses, namely, filtering, downsampling, normalization, and quantization.

- 1) *Filtering:* A second-order Butterworth low-pass filter with a 2-Hz cutoff frequency is applied to reduce the influence of noise. Because a vehicle is a relatively large mass, the behaviors of concern in our system generally appear in the range of the low-frequency part.
- 2) Downsampling: The input data with a high sampling rate have to be downsampled to reduce computational load. Both the data from the accelerometer and the camera are downsampled to 5 Hz. If we take the accelerometer as an example, we see that it provides three-axis acceleration at a sampling rate of 50 Hz. The mean is calculated every ten samples as a new value. This implies that the original sampling rate of 50 Hz is downsampled to a rate of 5 Hz. The downsampled data are gathered through a sliding window. We assume that the normal behavior of a vehicle

is presented within 5 s, which means that there are 250 samples of raw data and 25 samples after downsampling. Therefore, the length of the sliding window is set to 25 to acquire sufficient data.

- 3) Normalization: Normalization, which is performed to standardize the range of independent variables or features of data, is applied. Since the range of values of raw data widely varies in many machine learning algorithms, objective functions will not properly work without normalization. The selected HMM method is no exception. The minimum and the maximum are selected from the 25 samples in the window, and these two values are used to scale the data range into [0, 1]. In particular, instead of such commonly used normalization methods as min-max normalization and mean-variance normalization, a modified version of min-max normalization, which is called variance-based min-max normalization, is used for the following reasons. Conventional min-max normalization preserves the trend of the data but loses the amplitude information, which is the key feature of driving events. For example, the min-max normalization results of normal driving and zigzag driving are quite similar, as shown in Fig. 4. Therefore, variance-based min-max normalization checks the variance to remap the range of normalization to preserve the amplitude information. If the data variance $< T_{\rm var}$, where $T_{\rm var}$ is the maximum value of the statistics from the data of normal driving, the range is remapped to [0, 0.5] instead of [0, 1], as shown in Fig. 5.
- 4) Quantization: Vector quantization is introduced to generate distinct observation symbols for HMM usage. Vector quantization works by encoding values from a multidimensional vector space into a finite set of values from a discrete subspace of a lower dimension. The dimension of quantized data represents the number of symbols used in an HMM, and it affects HMM complexity. Quantized data with a low dimension may eliminate the variation of the data, whereas data with a high dimension increase computational complexity. By observing the recorded data, we chose a dimension of 10 to preserve the trend of the data. This means that the normalized value in [0, 1] is quantized with an integer in [1] and [10].

B. Behavior Recognition

After data preprocessing is accomplished, the HMM is applied to recognize the driving events. An HMM is a probabilistic tool for time-series data recognition. Owing to its stochastic nature, the HMM has been successfully used in a wide range of applications in the area of pattern recognition, particularly in speech recognition. Theories of HMMs were introduced by Baum and Petrie in the late 1960s [40]. Here, only the basic concepts are introduced. Detailed tutorials on HMMs can be found in [41].

An HMM is characterized by the following elements:

- *N*, the number of states in the HMM model;
- *M*, the number of distinct observation symbols per state;



Fig. 4. Results of min-max normalization. (a) Normal driving. (b) Zigzag driving.



Fig. 5. Results of variance-based min-max normalization. (a) Normal driving. (b) Zigzag driving.

- The state transition probability distribution $A = \{a_{ij}\}$, where $a_{ij} = P[q_{t+1} = S_j | q_t = S_i], 1 \le i, j \le N$;
- The observation symbol probability distribution in state j, $B = \{b_i(k)\}$, where

 $b_j(k) = P[v_k \text{ at } t | q_t = S_j], \qquad 1 \le j \le N, \ 1 \le k \le M$

• The initial probability distribution $\pi = \{\pi_i\}$, where

$$\pi_i = P[q_1 = S_i], \qquad 1 \le i \le N.$$

Therefore, an HMM λ could be specified as $\lambda = (N, M, A, B, \pi)$. The observation probability of sequence O is $P(O|\lambda)$. Moreover, there are different types of HMMs according to the limitation of state probability matrix A. The architecture of HMMs adopted in the algorithm is the so-called left-to-right model. The left-to-right model always starts from the first state, and transitions are only allowed toward the right state or the same state. The left-to-right model is better than the general model at performing dynamic pattern recognition, such as speech recognition, gesture recognition, and signature verification, because it lays a greater emphasis on the contextual relationship between states.

For each driving event, an HMM is constructed. As previously mentioned, a left-to-right HMM was adopted. The training phase is carried out using the Baum–Welch reestimation method. A larger number of symbols would reduce the quantization error but would, at the same time, reduce the recognition rate and increase computation complexity. On the other hand, an HMM with a large number of states leads to a larger error rate. Therefore, the number of states and symbols in each model is tested to find the optimal result. In all cases, except for ND, the number of states is five; for ND, the number of states is two. The number of all symbols is ten.

The driving event recognition procedure is shown in Fig. 6. Given an observation sequence, the observation probabilities of each HMM model are calculated by the forward algorithm. The model with the highest probability is then selected as the recognized maneuver. In other words, for one observation sequence set, one of the longitudinal behaviors is output as the result, such as DEC. The same is true for the lateral and carfollowing behaviors.



Fig. 6. HMM recognizers for three categories of driving behaviors.

IV. DANGER-LEVEL REASONING

The major task in the second stage is to combine the effect of each driving event to infer the danger level using an FIS. The degree of each driving event is estimated as a quantifiable indicator that represents a more explicit description. In addition, a hierarchical decision strategy is also presented to improve the efficiency of the fuzzy rule selection by using the recognized results obtained from the first stage as conditions. This is discussed in more detail in the following sections.

A. Degree of Driving Events

In the first stage, we analyzed the longitudinal, lateral, and car-following behaviors and recognized the occurrence of driving events in terms of these three aspects. In the second stage, the intent is to further estimate the danger level using the information on driving events. Each driving event itself implies a certain degree of risk; however, we have a rough criterion if we infer the danger level using the simple occurrence of each event. Take the case of acceleration, for example. One car accelerates at 10 km/h in 5 s, whereas another car accelerates at 30 km/h in 5 s. Evidently, the latter car is more dangerous. Therefore,

TABLE II Representative Parameters of Driving Events

Driving events	Representative parameters	Unit
Acceleration	ΔV	km/h
Deceleration	ΔV	km/h
Changing left	V _{lateral}	m/s
Changing right	V _{lateral}	m/s
Zigzag driving	$\sigma^2_{acc_y}$	$(m/s^2)^2$
Approaching the car in front	d	m

TABLE III LINGUISTIC VARIABLES OF FUZZY INPUT AND OUTPUT

Input/Output	Linguistic variables
Acceleration/Deceleration	AD
Zigzag driving	ZD
Lane changing	LC
Approaching the car in front	AFC
Speed	SP
Danger level	DL

considering only the occurrence of an event is not enough; the degree of that event must be also taken into account.

Table II lists the representative parameters of different degrees of each driving event. For acceleration and deceleration, variation in velocity, i.e., Δv , was chosen to reflect the intensity of the event. For lane changing, it is more dangerous if the vehicle changes lanes more rapidly. Thus, lateral velocity V_{lateral} is calculated to indicate the degree of changing to the left and right. For zigzag driving, the variance of the *y*-axis acceleration was selected, because the greater the degree of zigzagging, the larger the variance is. The degree of approaching the car in front is easy to imagine; the shorter the distance to the car in front, the greater the danger is.

B. Design of Fuzzy Danger-Level Model

After the driving events and their degrees have been recognized and estimated, an FIS is introduced to provide a reasonable danger level. Fuzzy systems are based on fuzzy logic, which was first developed by Zadeh in the mid-1960s to represent uncertain and imprecise knowledge [42]. The behavior of systems that are too complex or that cannot be easily analyzed mathematically can be described in an approximate but effective manner. A typical fuzzy logic model consists of four components, namely, fuzzification, a fuzzy inference engine, defuzzification, and a fuzzy rule base. The FIS adopts minimum operator as the connective AND, maximum as the aggregation method, and center of area as the defuzzification method.

The inputs of the FIS for the danger-level model include the variation in velocity, the variance of the y-axis acceleration, the lateral velocity, and the distance. The current speed of the host vehicle is also included because it also involves a different degree of risk. The output is the danger level. The linguistic variables of all the inputs and output are listed in Table III, and the membership functions are shown in Fig. 7.



Fig. 7. Membership functions. (a) Input membership functions for SP. (b) Input membership functions for AD. (c) Input membership functions for AFC. (d) Input membership functions for ZD. (e) Input membership functions for LC. (f) Output membership functions for DL.

C. Hierarchical Decision Strategy of Fuzzy Rule Selection

After the fuzzy sets and membership functions are set, the rule base for the model contains 72 fuzzy IF-THEN rules. If there are too many rules, however, this increases the computational load and leads to an imprecise understanding of the relationship between the behaviors and the danger level. To overcome these shortcomings, a hierarchical decision strategy was designed to decrease the number of fuzzy rules. Such a strategy involves three concepts. First, we know that some behaviors are mutually exclusive, i.e., they cannot simultaneously occur. For example, when a vehicle is zigzagging in a lane, the behavior of lane changing cannot happen at the same time. Therefore, if we know that the current behavior is zigzag driving, the parameters corresponding to lane changing can be ignored, thus reducing the number of combinations. Second, if the behavior itself does not exist, its effect does not need to be considered. For example, the behavior of a host vehicle approaching the car in front of it can only be observed when there actually is a car in front of it. Third, if the parameters of the corresponding behaviors are the same, they can be merged by using the same membership function. For example, the representative parameters of acceleration and deceleration are both variations of longitudinal acceleration. Therefore, we can use the one linguistic variable AD to denote their effects without using two separate inputs. In addition, the influence of acceleration or deceleration can be identified according to the sign of the parameter. In a similar sense, we also use another linguistic variable, i.e., LC, to indicate the degrees of changing



Fig. 8. Hierarchical decision flow.

TABLE IV RULE BASE OF HIGH-RISK SET

	AFC	NEAR			FAR			
ZD	SP AD	LOW	MEDIUM	HIGH	LOW	MEDIUM	HIGH	
	PB	Н	Н	Н	L	М	М	
LOW	ZO	М	Н	Н	L	М	М	
	NB	L	М	М	L	L	L	
	PB	Н	Н	Н	L	М	М	
HIGH	ZO	М	Н	Н	L	М	М	
	NB	М	М	Н	L	М	М	

TABLE V Rule Base of Medium-Risk Set

LC	AD SP	LOW	MEDIUM	HIGH
	PB	М	М	М
LOW	ZO	L	М	М
	NB	L	L	L
	PB	М	М	Н
HIGH	ZO	L	М	М
	NB	L	L	Н

left and changing right. By applying these three conditions, the fuzzy rules can be reduced by following the hierarchical decision flow shown in Fig. 8.

If the host car is approaching the car in front of it, this is undoubtedly a behavior belonging to the class of high risk. If the host car keeps a safe distance and stays in its lane, the risk level is low. Other cases are in the class of medium risk. Tables IV-VI show the fuzzy rule bases of high-risk, mediumrisk, and low-risk sets, respectively. For purposes of clarity, we use the following abbreviations: L = low; M = medium; H =high; PB = positive big; ZO = zero; and NB = negative big. According to the different conditions, one of the three rule bases is chosen, and one of the rules is triggered. Because there are different inputs in each rule base, this leads to a different number of combinations of rules. Obviously, the high-risk set generates 36 rules, whereas both the medium-risk and low-risk sets generate 18 rules. Although there is a total of 72 rules in the fuzzy rule base, the system has to first decide which risk set to choose and then trigger the rules from the selected risk set. Therefore, the maximum number of rules used is 36, and the average number is (36 + 18 + 18)/3 = 24.

TABLE VI Rule Base of Low-Risk Set

ZD	AD	LOW	MEDIUM	HIGH
	PB	L	L	М
LOW	ZO	L	L	М
	NB	L	L	L
	PB	L	L	М
HIGH	ZO	L	L	М
	NB	L	L	М



Fig. 9. System prototype.

V. EXPERIMENTAL RESULTS

This section addresses the implementation details and the experimental results. All functionalities were successfully implemented and installed in our experimental car. The camera was mounted on the front windshield, and the accelerometer and the Bluetooth GPS receiver were attached on top of the dashboard. A prototype was implemented on a platform based on a TI DM3730 processor, which has a 1-GHz Advanced RISC Machine core and an 800-MHz digital signal processor core, as shown in Fig. 9. The captured image size was 320×240 (QVGA), and the processing performance achieved 30 frames/s for lane marking and vehicle detection.

All data for training and testing were recorded with the experimental car in the environment of the freeway, highway, and various suburban routes in Taiwan. The freeway and highway were used to record the data at high speed, whereas the suburban routes were used to record data at low speed. Three test drivers were instructed to execute all driving events under safe conditions. The recorded data were manually selected for the training and testing of the HMMs of all driving events. The *x*-axis acceleration was used for longitudinal behaviors, the lane bias for lateral behaviors, and the distance to the car in front for car-following behaviors. The selected data segmentations of each driving event are shown in Fig. 10.

The results for driving event recognition are shown in Table VII. The first row lists the names and the number of evaluated driving events. For example, (ND, 193) means that 193 sequences of normal driving were evaluated. Each column in the table represents the number of occurrences of each driving event that was recognized as being other driving events. In an ideal case, all cells in the table should be zero, except for the diagonal values, which should be equal to the number of corresponding driving events. 30% of the data of each evaluated sequence were used to train each HMM, whereas the remainder were used to evaluate the detection ratio of recognition.



Fig. 10. Data sequences for training and testing. (a) Acceleration. (b) Deceleration. (c) Changing left. (d) Changing right. (e) Zigzag driving. (f) Approaching the car in front.

TABLE VII CONFUSION MATRIX OF DRIVING EVENT RECOGNITION

	Evaluated sequences								
Recognized	ND,	ACC,	DEC,	CL,	CR,	ZD,	AFC,		
as	193	115	110	162	160	154	102		
ND	193	2	0	0	0	2	0		
ACC	0	113	0	0	0	0	0		
DEC	0	0	110	0	0	0	0		
CL	0	0	0	161	0	0	0		
CR	0	0	0	0	158	0	0		
ZD	0	0	0	1	2	152	0		
AFC	0	0	0	0	0	0	102		

It can be seen that ND, DEC, and AFC were perfectly discriminated by the classifier. Although the behavior of acceleration and deceleration is similar, ACC may have been misrecognized as ND, owing to the variation of acceleration as compared with DEC, which is not so obvious. This can be also understood from the different delaying effects of the throttle and brakes. CL was misrecognized as ZD because the vehicle changed to the left lane too slowly and the driver made a slight turn to the right. The reason for the misrecognized case of CR is also similar. In the error case of ZD, it was misrecognized as ND. Because ZD and ND primarily show a similar trend,

the data of ZD with a smaller variation are confused with that of ND.

The HMM is the recognition approach chosen for the first stage. There are four highly related approaches to recognition in driving event detection. An SVM and logistic regression are used to detect the driver's state [19]. An ANN is used to detect driver fatigue [26]. A Bayesian filter is used to detect the driver's behavior at crossroads [35]. All of these are well-known machine learning methods that have been used to solve many kinds of problems.

A comparison of the results obtained for the detection ratio is shown in Fig. 11. These results prove that the proposed design produces better performance, because it resulted in higher detection ratios than the other four approaches.

In particular, for the cases of ZD and AFC, the HMM had a higher detection ratio than the other methods. For SVM, ANN, Bayesian filter, and logistic regression, the statistical characteristics, e.g., mean, variance, minimum, and maximum, are composed as a feature vector that is used to identify the driving events. Therefore, if these statistical characteristics cannot effectively distinguish the data of recognized driving events, the results indicate poor performance. For the HMM, however, the trend of the data is the main feature used for recognition, and



Fig. 11. Comparison of the detection ratio for recognition results.

 TABLE
 VIII

 COMPARISON OF FALSE-ALARM RATES (BY PERCENTAGE)

	ND	ACC	DEC	CL	CR	ZD	AFC	Avg.
HMM	2.03	0	0	0	0	1.93	0	0.57
SVM	11.14	27.74	1.92	2.45	9.77	8.27	7.61	9.84
ANN	7.10	2.73	6.14	0.63	1.26	6.83	16.36	5.86
BF	14.64	17.77	1.92	2.44	9.83	5.07	8.33	8.57
LR	14.66	26.97	3.85	1.24	6.54	6.08	7.61	9.56

the proposed variance-based min-max normalization also increases the level of discrimination for different driving events. These characteristics resulted in better performance for ZD and AFC.

In addition to a high detection ratio, a low false-alarm rate is also significant. Table VIII lists the comparative results of the false-alarm rates for the proposed approach and for the other four approaches. Except for ND and ZD, the other driving events produced no false alarms. In Table VII, we can see that the false alarms for ZD are due to the misrecognition of CL and CR and that those for ND are due to the misrecognition of ACC and ZD. The average false-alarm rate was only 0.57%, which is much lower than that of the other approaches. We would recall here that the recognition process involves comparing the probability of each model and choosing the model with the highest probability as being the recognized driving event. Therefore, the difference between the highest probability and the second highest probability reflects the robustness of recognition. The average ratio of the highest probability and the second highest probability of recognized driving events in all successful identifications was 2.17. This means that the highest probability value is twice that of the second highest value, which indicates that the recognition process is quite robust.

Fig. 12 shows a safe driving course, on which the driver maintained a smooth pace and a steady path, and there were no cars in the front. The longitudinal, lateral, and car-following behaviors were all recognized as representing normal driving. The details of the driving data were as follows: The speed was 95 km/h, and the variation of speed was -8.43 km/h. The variance of the *y*-axis acceleration was 0.107 (m/s²)².

Owing to the hierarchical decision strategy of danger-level reasoning, the rule base of low risk was invoked. Two rules were therefore activated.



Fig. 12. Safe driving course.



Fig. 13. Dangerous driving situation.

- 1) If ZD is LOW and AD is ZO and SP is HIGH, then DL is MEDIUM.
- 2) If ZD is LOW and AD is NB and SP is HIGH, then DL is LOW.

The evaluated danger level was 40.2, which represents a safe driving situation. Therefore, the safety indicator on the upper right-hand corner in Fig. 12 shows a green sign. This way, a driver can understand his/her situation at a glance.

Fig. 13 shows a relatively more dangerous situation. The driver speeded up and came close to the car in front. This is definitely dangerous driving. The details of the driving data were as follows: The speed was 83 km/h, and the variation of speed was 23.4 km/h. The variance of the *y*-axis acceleration was $0.109 \text{ (m/s}^2)^2$. The distance to the car in front was 10.5 m.

Because of the existing AFC, the rule base of high risk was invoked. These two rules were therefore activated.

- 1) If AFC is NEAR and ZD is LOW and AD is PB and SP is HIGH, then DL is HIGH.
- 2) If *AFC* is *FAR* and *ZD* is *LOW* and *AD* is *PB* and *SP* is *HIGH*, then *DL* is *MEDIUM*.

The evaluated danger level was 81.3, and the safety indicator showed a red sign. The driver would immediately know that he/she is in danger and must slow down to avoid an accident.

The degree of danger level is determined by the fuzzy rules. However, for the inferred danger level, e.g., the system shows a yellow indicator, some drivers might think that the situation is much dangerous, whereas some might feel just appropriate. We believe that this circumstance exists due to different characteristics in different persons. Therefore, our evaluation strategy focuses on examining the accuracy of the occurrence of medium- and high-danger degrees. Three test drivers were instructed to drive on the freeway for about 30 min. Two kinds of tests were performed. The first kind of test is to watch out for circumstances and drive in a normal and safe way. The inferred danger levels were recorded, and we expected that all of them revealed the low danger level. The results with medium or high danger levels were represented as false alarms. The second kind of test is to behave dangerous actions consciously and carefully, e.g., changing lane rapidly. The inferred danger levels were manually marked every time the driver started doing the dangerous action. We expected that none of the marked danger levels revealed the low danger level. The results with low danger levels were represented as false alarms. The first and second types of tests were performed by three drivers for 10 times in each case, whereas 20 dangerous actions were behaved during the second test. The results showed that there are no false alarms in both tests, which means that the performance of the second stage in the proposed framework is applicable.

VI. CONCLUSION

This study has proposed a two-stage reasoning-based framework that uses driving event recognition to provide drivers with an intuitive danger-level indicator for driving safety. Except for the information on driving events, a single index to evaluate a driver's performance is very useful, because it is convenient to know drivers' status on their routes. People could more easily clarify the cause of an accident or know the stability of a driver by examining the occurring driving events according to danger levels. To overcome the issue of labeling that arises from the uncertain definition of a dangerous pattern, the danger level is inferred by an FIS, using information based on recognized driving events. Our system not only informs drivers of their current safety status but offers information regarding the possible behaviors that led to this status as well. To reduce the complexity of the FIS, a hierarchical decision strategy was proposed, which can significantly decrease the number of rules. The developed functions were fully implemented on an embedded platform and were then tested in a real environment. All HMM training models for seven driving events achieved an average detection ratio of approximately 99%. Two types of tests were designed to evaluate the inferred results, and the

results signified that the danger-level indicator could provide suitable notifications for driving safety monitoring. However, the design of danger-level inference is insufficient because the ground truth of a danger level is not available in this study. Although the evaluation shows that the results are functional for notification, we cannot conclude that the inferred danger level is accurate. In the future, a valid approach to directly or indirectly estimate the value of danger will be investigated. Then, the methods for automatic rule generation and structure optimization are possible to apply in the inference stage in our framework to acquire more comprehensive inferred results.

REFERENCES

- C. M. Feng, "Road accidents profile in Taiwan," J. Int. Assoc. Traffic Safety Sci., vol. 31, no. 1, pp. 138–139, Feb. 2007.
- [2] T. Ehlgen, T. Pajdla, and D. Ammon, "Eliminating blind spots for assisted driving," *IEEE Trans. Intell. Transp. Syst.*, vol. 9, no. 4, pp. 657–665, Dec. 2008.
- [3] 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. Intell. Transp. Syst.*, vol. 7, no. 3, pp. 293– 308, Sep. 2006.
- [4] V. Kastrinaki, M. Zervakis, and K. Kalaitzakis, "A survey of video processing techniques for traffic applications," *Image Vis. Comput.*, vol. 21, no. 4, pp. 359–381, Apr. 2003.
- [5] Y. Dong, Z. Hu, K. Uchimura, and N. Murayama, "Driver inattention monitoring system for intelligent vehicles: A review," *IEEE Trans. Intell. Transp. Syst.*, vol. 12, no. 2, pp. 596–614, Jun. 2011.
 [6] A. Doshi and M. M. Trivedi, "Tactical driver behavior prediction and
- [6] A. Doshi and M. M. Trivedi, "Tactical driver behavior prediction and intent inference: A review," in *Proc. IEEE Intell. Transp. Syst.*, 2011, pp. 1892–1897.
- [7] B. Atsumi, "Evaluation of mental condition on drivers by analysis of heart rate variability: Measurement of mental stress and drowsiness by indexes of autonomic nervous system," *JSAE Rev.*, vol. 16, no. 1, p. 110, Jan. 1995.
- [8] K. Jiao, Z. Li, M. Chen, C. Wang, and S. Qi, "Effect of different vibration frequencies on heart rate variability and driving fatigue in healthy drivers," *Int. Arch. Occup. Environ. Health*, vol. 77, no. 3, pp. 205–212, Apr. 2004.
- [9] B. T. Jap, S. Lal, P. Fischer, and E. Bekiaris, "Using EEG spectral components to assess algorithms for detecting fatigue," *Exp. Syst. Appl.*, vol. 36, no. 2, pp. 2352–2359, Mar. 2009.
- [10] M. V. M. Yeo, X. Li, K. Shen, and E. P. V. Wilder-Smith, "Can SVM be used for automatic EEG detection of drowsiness during car driving?" *Safety Sci.*, vol. 47, no. 1, pp. 115–124, Jan. 2009.
- [11] K. Q. Shen, X. P. Li, C. J. Ong, S. Y. Shao, and E. P. V. Wilder-Smith, "EEG-based mental fatigue measurement using multi-class support vector machines with confidence estimate," *Clin. Neurophysiol.*, vol. 119, no. 7, pp. 1524–1533, Jul. 2008.
- [12] C. T. Lin, R. C. Wu, S. F. Liang, W. H. Chao, Y. J. Chen, and T. P. Jung, "EEG-based drowsiness estimation for safety driving using independent component analysis," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 52, no. 12, pp. 2726–2738, Dec. 2005.
- [13] C. P. Chua, G. McDarby, and C. Heneghan, "Combined electrocardiogram and photoplethysmogram measurements as an indicator of objective sleepiness," *Physiol. Meas.*, vol. 29, no. 8, pp. 857–868, Aug. 2008.
- [14] S. K. L. Lal and A. Craig, "A critical review of the psychophysiology of driver fatigue," *Biol. Psychol.*, vol. 55, no. 3, pp. 173–194, Feb. 2001.
- [15] G. Yang, Y. Lin, and P. Bhattacharya, "A driver fatigue recognition model based on information fusion and dynamic Bayesian network," *Inf. Sci.*, vol. 180, no. 10, pp. 1942–1954, May 2010.
- [16] L. M. Bergasa, J. Nuevo, M. A. Sotelo, R. Barea, and M. E. Lopez, "Realtime system for monitoring driver vigilance," *IEEE Trans. Intell. Transp. Syst.*, vol. 7, no. 1, pp. 63–77, Mar. 2006.
- [17] Q. Ji and X. Yang, "Real-time eye, gaze, and face pose tracking for monitoring driver vigilance," *Real-Time Imag.*, vol. 8, no. 5, pp. 357–377, Oct. 2002.
- [18] Q. Ji and X. Yang, "Real time visual cues extraction for monitoring driver vigilance," in *Proc. Comput. Vis. Syst.*, 2001, vol. 2095, pp. 107–124.
- [19] L. Yulan, M. L. Reyes, and J. D. Lee, "Real-time detection of driver cognitive distraction using support vector machines," *IEEE Trans. Intell. Transp. Syst.*, vol. 8, no. 2, pp. 340–350, Jun. 2007.

- [20] A. B. Albu, B. Widsten, W. Tiange, J. Lan, and J. Mah, "A computer vision-based system for real-time detection of sleep onset in fatigued drivers," in *Proc. IEEE Intell. Veh. Symp.*, 2008, pp. 25–30.
- [21] J. Qiang, Z. Zhiwei, and P. Lan, "Real-time nonintrusive monitoring and prediction of driver fatigue," *IEEE Trans.Veh. Technol.*, vol. 53, no. 4, pp. 1052–1068, Jul. 2004.
- [22] H. Ueno, M. Kaneda, and M. Tsukino, "Development of drowsiness detection system," in *Proc. Veh. Navig. Inf. Syst. Conf.*, 1994, pp. 15–20.
- [23] F. Friedrichs and B. Yang, "Drowsiness monitoring by steering and lane data based features under real driving conditions," in *Proc. Eur. Signal Process. Conf.*, 2010, pp. 209–213.
- [24] Y. J. Zhong, L. P. Du, K. Zhang, and X. H. Sun, "Localized energy study for analyzing driver fatigue state based on wavelet analysis," in *Proc. Wavelet Anal. Pattern Recognit.*, 2007, vol. 4, pp. 1843–1846.
- [25] K. Torkkola, N. Massey, and C. Wood, "Driver inattention detection through intelligent analysis of readily available sensors," in *Proc. IEEE Intell. Transp. Syst.*, 2004, pp. 326–331.
- [26] A. Eskandarian and A. Mortazavi, "Evaluation of a smart algorithm for commercial vehicle driver drowsiness detection," in *Proc. IEEE Intell. Veh. Symp.*, 2007, pp. 553–559.
- [27] S. Zehang, G. Bebis, and R. Miller, "On-road vehicle detection: A review," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 5, pp. 694–711, May 2006.
- [28] J. C. McCall and M. M. Trivedi, "Video-based lane estimation and tracking for driver assistance: Survey, system, and evaluation," *IEEE Trans. Intell. Transp. Syst.*, vol. 7, no. 1, pp. 20–37, Mar. 2006.
- [29] Y. Liu, Z. Han, Z. Zhang, and J. Jiao, "Video-based hazardous event detection in the driving environment," in *Proc. Eng. Manage.*, 2009, pp. 736–740.
- [30] K. Inata, P. Raksincharoensak, and M. Nagai, "Driver behavior modeling based on database of personal mobility driving in urban area," in *Proc. Int. Conf. Control, Autom. Syst.*, 2008, pp. 2902–2907.
- [31] Y. R. Huang, "Real time dangerous driving status detection," in *Proc. Int. Congr. Image Signal Process.*, 2011, vol. 1, pp. 414–418.
- [32] Y. Zhou, W. Xu, H. Ning, Y. Gong, and T. Huang, "Detecting unsafe driving patterns using discriminative learning," in *Proc. IEEE Int. Conf. Multimedia Expo.*, 2007, pp. 1431–1434.
- [33] H. Ning, W. Xu, Y. Zhou, Y. Gong, and T. Huang, "A general framework to detect unsafe system states from multisensor data stream," *IEEE Trans. Intell. Transp. Syst.*, vol. 11, no. 1, pp. 4–15, Mar. 2010.
- [34] J. Wang, S. Zhu, and Y. Gong, "Driving safety monitoring using semisupervised learning on time series data," *IEEE Trans. Intell. Transp. Syst.*, vol. 11, no. 3, pp. 728–737, Sep. 2010.
- [35] G. S. Aoude, V. R. Desaraju, L. H. Stephens, and J. P. How, "Driver behavior classification at intersections and validation on large naturalistic data set," *IEEE Trans. Intell. Transp. Syst.*, vol. 13, no. 2, pp. 724–736, Jun. 2012.
- [36] S. Sekizawa, S. Inagaki, T. Suzuki, S. Hayakawa, N. Tsuchida, T. Tsuda, and H. Fujinami, "Modeling and recognition of driving behavior based on stochastic switched ARX model," *IEEE Trans. Intell. Transp. Syst.*, vol. 8, no. 4, pp. 593–606, Dec. 2007.
- [37] A. Sathyanarayana, S. Nageswaren, H. Ghasemzadeh, R. Jafari, and J. H. Hansen, "Body sensor networks for driver distraction identification," in *Proc. IEEE ICVES*, 2008, pp. 120–125.
- [38] C. J. Chen, H. Y. Peng, B. F. Wu, and Y. H. Chen, "A real-time driving assistance and surveillance system," J. Inf. Sci. Eng., vol. 25, no. 5, pp. 1501–1523, Sep. 2009.
- [39] B. F. Wu, Y. H. Chen, C. C. Kao, Y. F. Li, and C. J. Chen, "A vision-based collision warning system by surrounding vehicles detection," *KSII Trans. Internet Inf. Syst.*, vol. 6, no. 4, pp. 1203–1222, Apr. 2012.
- [40] L. E. Baum and T. Petrie, "Statistical inference for probabilistic functions of finite state Markov chains," *Ann. Math. Stat.*, vol. 37, no. 6, pp. 1554– 1563, Dec. 1966.
- [41] L. R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proc. IEEE*, vol. 77, no. 2, pp. 257–286, Feb. 1989.
- [42] L. A. Zadeh, "Fuzzy sets," *Inform. Control*, vol. 8, no. 3, pp. 338–353, Jun. 1965.



Bing-Fei Wu (S'89–M'92–SM'02–F'12) received the B.S. and M.S. degrees in control engineering from National Chiao Tung University (NCTU), Hsinchu, Taiwan, in 1981 and 1983, respectively, and the Ph.D. degree in electrical engineering from the University of Southern California, Los Angeles, CA, USA, in 1992.

Since 1992, he has been with the Department of Electrical Engineering, NCTU, where he was promoted to Professor in 1998 and Distinguished Professor in 2010. Since 2011, he has served as the

Director of the Institute of Electrical and Control Engineering, NCTU. His current research interests include image recognition, vehicle driving safety, intelligent control, intelligent transportation systems, multimedia signal analysis, and embedded systems.

Dr. Wu was elevated as a Fellow of the IEEE for his contributions to intelligent transportation and multimedia systems. Since 2011, he has been a Fellow of the Institution of Engineering and Technology. In 2003, he founded and served as the Chair of the IEEE Systems, Man, and Cybernetics (SMC) Society Taipei Chapter. In 2011, he was elected as the Chair of the Technical Committee on Intelligent Transportation Systems of the IEEE SMC Society. He has received many research honors, including the Outstanding Research Award from the Pan Wen-Yuan Foundation in 2012; the Best Technology Transfer Award from the National Science Council, Taiwan, in 2008; and the Outstanding Information Technology Elite Award in 2003.



Ying-Han Chen was born in Tainan, Taiwan, in 1981. He received the B.S. and M.S. degrees in electrical engineering from National Central University, Jhongli, Taiwan, in 2003 and 2006, respectively. He is currently working toward the Ph.D. degree in electrical control engineering with National Chiao Tung University, Hsinchu, Taiwan.

His research interests include computer networks, embedded systems, and digital signal processing.





His research interests include signal and image processing and embedded systems.



Yen-Feng Li was born in Taitung, Taiwan, in 1976. He received the B.S. degree in electronic engineering from National Taiwan University of Science and Technology, Taipei, Taiwan, in 2007 and the M.S. degree in electronic engineering from National Chin-Yi University of Technology, Taichung, Taiwan, in 2009.

He is currently with CSSP Inc., Hsinchu, Taiwan. His research interests include image processing, embedded systems, and intelligent transportation systems.