The DRIVE-SAFE Project: Signal Processing and Advanced Information Technologies for Improving Driving Prudence and Accidents

Aytül Erçil [(1)], Levent Güvenç [(2)], Hüseyin Abut [(3)], Engin Erzin[(4)], Ali Güvenç Göktan [(5)]

[(1)] Sabancı University, [(2)] Istanbul Technical University, [(3)] San Diego State University, [(4)] Koç University, [(5)] OTAM Automotive Research Center

Abstract

In this paper, we will talk about the Drivesafe project whose aim is creating conditions for prudent driving on highways and roadways with the purposes of reducing accidents caused by driver behavior. To achieve these primary goals, critical data is being collected from multimodal sensors (such as cameras, microphones, and other sensors) to build a unique databank on driver behavior. We are developing system and technologies for analyzing the data and automatically determining potentially dangerous situations (such as driver fatigue, distraction, etc.). Based on the findings from these studies, we will propose systems for warning the drivers and taking other precautionary measures to avoid accidents once a dangerous situation is detected. In order to address these issues a national consortium has been formed including Automotive Research Center (OTAM), Koç University, Istanbul Technical University, Sabancı University, Ford A.Ş., Renault A.Ş., and Fiat A.Ş.

1 Introduction

Every year, more than three million accidents involving motor vehicles cause a total of 40000 deaths worldwide. Driver error has been blamed as the primary cause for approximately 80% of these accidents. According to the figures given by the US National Highway Traffic Safety Administration, driver fatigue has resulted in 240,000 fatalities in the U.S. alone. In addition, it is also reported that the sleep related accidents cost public and private sector over $46 billion every year.

According to the 2005 statistics prepared by the Traffic Education and Research Directorate of the Department of Security, Turkey, 3,215 people have died and 123,985 people have been injured in over 570,000 traffic accidents in highways and roads in Turkey. Primary economical loss due to these traffic accidents is estimated to be 651,166,236 USD. However, when the secondary and tertiary losses from loss of time, future income, the impact on the highways and the fuel costs included this future explodes to over 6.0 Billion US Dollars, which is a quite high percentage of the national GNP of the county.

In 2005, an academic and industry consortium called Drive Safe has been established to create conditions for prudent driving on highways and roadways with the purposes of reducing accidents resulting from abnormal driver behavior. The consortium consists of Sabancı University, Koç University, ITU Mekar Laboratories, OTAM Automotive Research Center, Ford A.Ş., Oyak Renault A.Ş. and Fiat A.Ş. Services of several experts both from Turkey and abroad have been acquired to assist the team. This multi-campus multi-disciplinary initiative is supported by the Turkish State Planning Organization and is partially supported by Japanese New Energy and Industrial Technology Development Organization (NEDO). In addition, all partners and a few sponsors are providing resources in terms of equipment, fuel, man-power, and services.

To achieve these primary goals, critical data will be collected from multimodal sensors including cameras, microphones, vehicular, and driver/driving related sensors to build a unique databank on driver behavior. The goal of the project is to develop systems and technologies for analyzing the data and automatically determining potentially dangerous situations, i.e., driver fatigue, distraction, drunk driving, etc. Based on the findings from these studies, the aim is to propose systems for warning the drivers and taking other precautionary measures to avoid accidents once a dangerous situation is detected.

Secondary objectives include but are not restricted to:

i. Personalization of vehicular chamber for improved safety and comfort.
ii. Secure, transparent and efficient communications in a hands-free environment.

There have been a lot of effort, especially within the European community to address the problems in road and traffic safety. For example, eSafety, the first pillar of the Intelligent Car Initiative, is a joint initiative of the European Commission, industry and other stakeholders and aims to accelerate the development, deployment and use of Intelligent Integrated Safety Systems, that use information and communication technologies in intelligent solutions, in order to increase road safety and reduce the number of accidents on Europe's roads.

Some of the subprojects that are under the umbrella of the eSafety project are: ADASE, AIDE, APROSYS, AWAKE, CarTALK, CHAUFFEUR 2, EASIS, EUCLUDE, ESCOPE, GST, HIGHWAY, HUMANIST.
http://europa.eu.int/information_society/activities/esafety/index_en.htm

The main goal of the ADASE 2 (Advanced Driver Assistance Systems in Europe) project is to introduce and implement active safety systems by harmonising and communicating active safety functions, identifying technological needs and focussing on essentials, preparing architectures, roadmaps and standards. http://www.adase2.net/ The AIDE (Adaptive Integrated Driver-vehicle InterfacE) project concentrates on human-machine interaction. http://www.aide-eu.org/

The APTOSYS (Integrated Project on Advanced Protection Systems) project has concentrated on passive safety systems. In this project different accidents have been studied and design implications for vehicles (passenger vehicles, heavy vehicles, motorcycles, etc.), and roads have been studied. http://www.aprosys.com/ The objective of the AWAKE project was to increase traffic safety by reducing the number and the consequences of traffic accidents caused by driver hypo-vigilance. The project has concentrated on detecting and diagnosing driver hypo-vigilance in real-time by fusing data from on-board driver monitoring sensors (eyelid behaviour and steering grip forces) and data regarding the driver’s behaviour (lane keeping performance) http://www.awake-eu.org/ The main goal of the EASIS (Electronic Architecture and System Engineering for Integrated Safety Systems) project is to develop powerful and highly dependable in-vehicle electronic architecture and appropriate development support for the realization of Integrated Safety Systems. http://www.easis-online.org/wEnglish/overview/index.shtml?navid=1

The SENSATION (Advanced Sensory Development for Attention, Stress, Vigilance and Sleep/Wakefulness Monitoring) project aims to explore a wide range of micro and nano sensor technologies, with the aim to achieve unobtrusive, cost-effective, real-time monitoring, detection and prediction of human physiological state in relation to wakefulness, fatigue and stress anytime, everywhere and for everybody. http://www.sensation-eu.org/

2 Data Collection

In this section we will discuss the various data collection activities in the project.

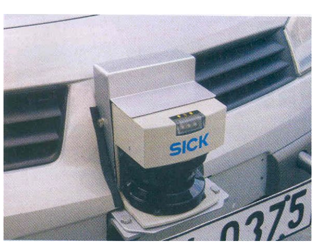
2.1 In-Car Multisensor Data Collection

A data collection vehicle was equipped with various sensors to collect data on driver behaviour under normal conditions. Figure 1 shows the test vehicle and Figure 2 shows some of the sensors in the vehicle.

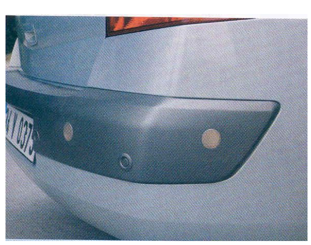
Figure 1: Data Collection vehicle

The sensors in the test vehicle include 3 day cameras (2 looking at the driver, one looking at the road), 3 night cameras, various microphones, an inertial measurement unit (IMU) measuring xyz accelerations and angular velocities, 2D laser scanner, gas and brake pedal pressure sensors, GPS receiver, EEG (electroencephalograph). Additional data like tire angular speeds, steering wheel position and speed, engine rotational speed, vehicle longitudinal speed, vehicle yaw rate, turn signal states, clutch pedal position switches, brake pedal position switch, idle gear state and rear gear switch are obtained from the CANbus. A Matlab graphical user interface for easy and interactive visualization of non-vision, non-sound data collected in each run was prepared and is being used.

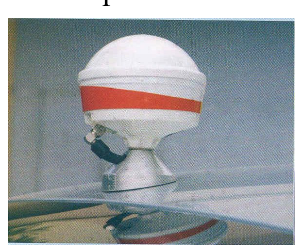
Driving signals cannot be employed for driver authorization, since authorization should be performed before driving starts. However, driving behavior signals can be used to verify the driving condition of an identified user in a safe driving scenario. Assuming that the driver has already been identified, the driving behavior signals can be used to verify whether the driver is alert, as opposed to being sleepy or drunk. In addition, they could be useful for forensic purposes to identify drivers in a stolen-and-found vehicle or after a crash when audio-visual sensors are not available or cannot be relied upon.



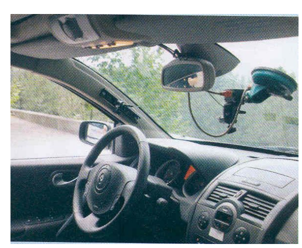
laser scanner



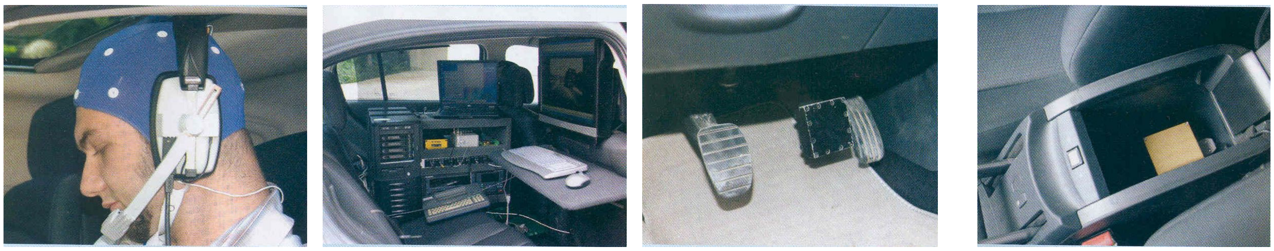
sonar sensors



GPS



Cameras



EEG and microphone | data collection center | Brake pedal sensor | 3 axis IMU

Figure 2. Various sensors in the data collection vehicle

Electroencephalography is the neurophysiologic measurement of the electrical activity of the brain from electrodes placed on the scalp. These signals showing brain activity will be used as ground truth for determining fatigue from other sensors.

Figure 3 shows the route to be used in the data collection experiments. This route will be used in modeling driver behaviour under normal conditions.

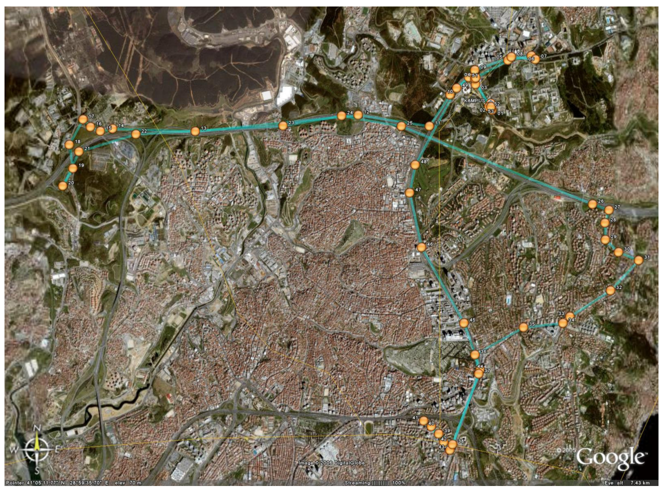
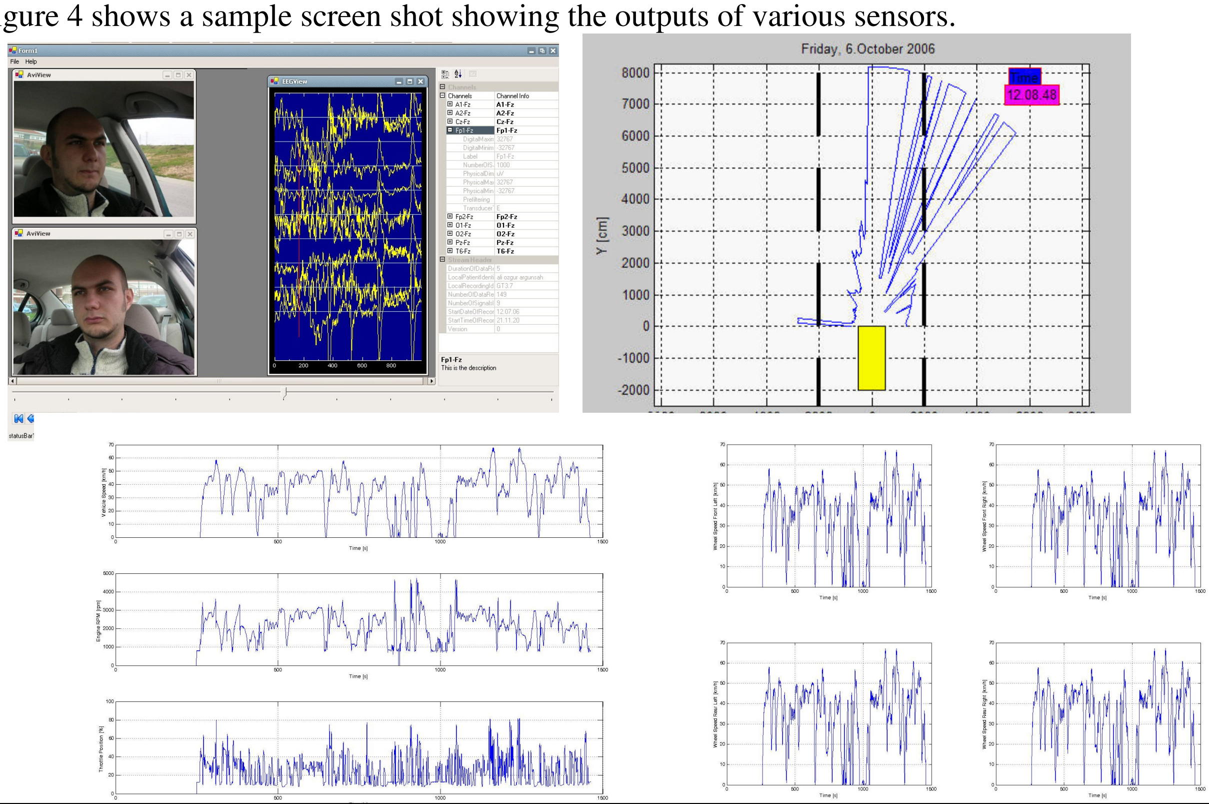


Figure 3. Experimental vehicle's route

Figure 4 shows a sample screen shot showing the outputs of various sensors.



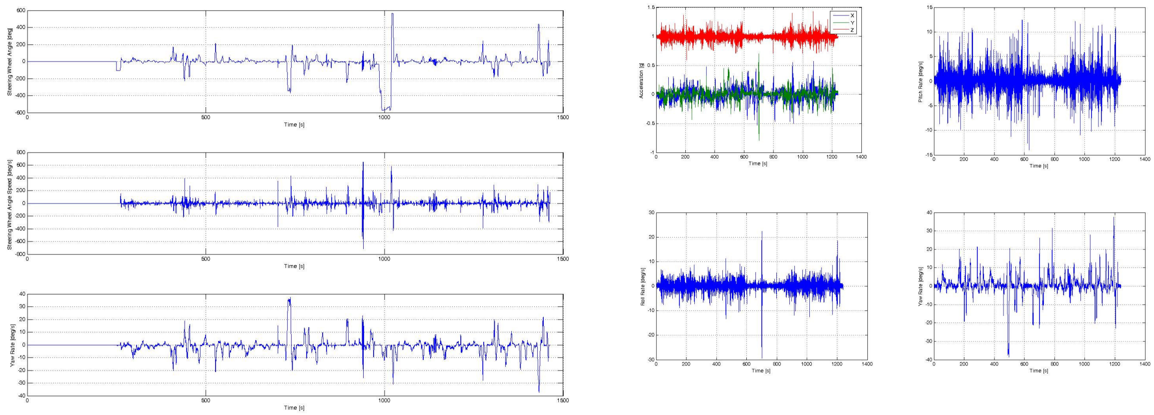


Figure 4. Sample screen shots showing the outputs of various sensors (cameras, EEG, laser sensor, CAN bus and IMU sensor)

2.2 Data Collection in a Simulator environment

Our experiments will focus on detecting driver fatigue after long hours of driving. Because of the dangers of carrying out these experiments in real driving conditions, data collection for these experiments will be carried out in a simulator environment. To check for the validity of the simulator data, the 3D model of the route that the test car will use has been made for the simulator. A separate route of 4-5 hours driving in various environments is designed to carry out fatigue experiments

2.3 Animation Data

Until sufficient data has been collected for fatigue modeling, an animation environment has been developed to generate fatigue conditions. For this purpose, muscle movements in the face are modeled, then video animations are synthesized for creating expressions related to fatigue. It is then possible to obtain different videos under different lighting conditions, different hair styles, etc.

How to quantitatively and completely describe the rich facial expressions is a long-term endeavor in both engineering and psychophysiology. Ekman's group [3] proposed a unified description method of expression: Facial Action Coding System (FACS). It totally has 71 primitive units, called Action Unit (AU). Based on them, any expression display can be represented by single AU or the AU combination.

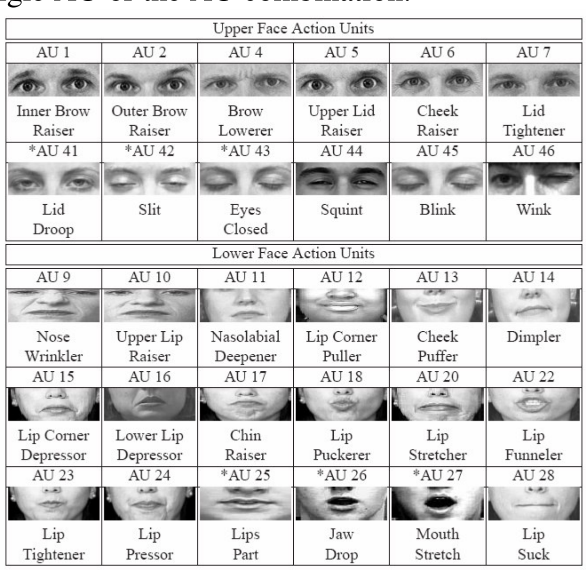


Figure 5. Sample action units

This system has been used in modeling muscle movements. Figure 6 shows some frames from action unit simulations.

Figure 6 (a) Some frames from the Action Unit 1 (inner brow raiser) video

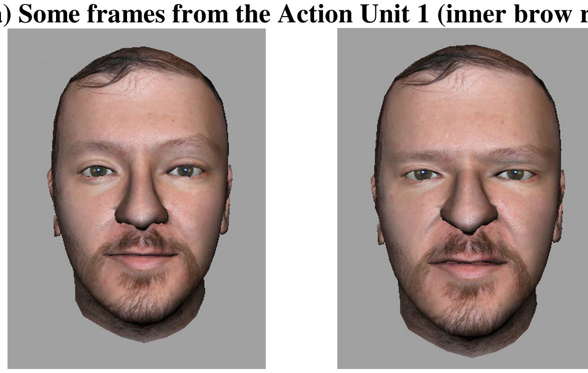


Figure 6 (b) Some frames from the Action Unit 9 (nose wrinkler) video

3 Person Recognition in Vehicles

It is expected that next generation human-vehicle interfaces will incorporate biometric person recognition using speech, video/image, analog driver behavior signals to provide more efficient and safe vehicle operation, as well as pervasive and secure communication in vehicles. Yet, technical and deployment limits exist in the ability of these systems to perform satisfactorily in real-world settings under adverse conditions such as in human-vehicle interaction. For instance, factors such as environmental noise, changes in acoustic and microphone conditions all significantly impact speaker recognition performance. Similarly, factors such as illumination variation, background variation, camera resolution and angle, and facial expressions contribute to performance loss of the visual person recognition. Biometric person recognition in vehicles is likely to be most challenging not only due to the challenges posed by the chamber, but also due to cost-economics. In this section we demonstrate that the required levels of accuracy for biometric person recognition in vehicles can be achieved by fusion of multiple modalities.

Personalization of vehicles will open up new services for personal and vehicle safety, including:

1. **Vehicle safety:** Determine whether the person behind the wheel is one of the authorized drivers.
2. **Safe driving**: Verify whether the driver is in normal condition, as opposed to being sleepy or drunk, for safe driving.
3. **Safety of the people and goods especially in commercial vehicles**
4. **Secure transactions:** The more connected we become, the more transactions we need to do anywhere we can, which includes inside a vehicle. These include travel planning and arrangements, m-banking[1], m-database access, and m-shopping, which all require varying levels of personal authentication.

Biometric person recognition has been extensively studied for more than two decades and technologies with varying degrees of success have been developed. However, most promising systems with respectable benchmark recognition rates, in controlled environment, have suffered significantly when it came to deployment in challenging environments such as cockpit of an airplane, or a moving vehicle chamber.

The task of recognizing people from in-vehicle video is difficult for the following reasons:
1. In vehicles, the subjects, especially the driver, are not expected to pose for the camera since their first priority is to operate the vehicle safely. Hence, there can be large illumination and pose variations. In addition, partial occlusions and disguise are common.

2. The spatial resolution and quality of video is usually low, and due to the acquisition conditions, the face image sizes are smaller (sometimes much smaller) than the assumed sizes in most existing still image based face recognition systems.

4 Experimental Results

With funding from Japanese government and industry, Itakura *et al.* in the Center for Acoustic Information Research (CIAIR) at Nagoya University have embarked on a mega project called "Construction and Analysis of the Multi-Layered In-Car Spoken Dialogue Corpus", where audio (12 channels), video (3 channels), analog driver behavior signals (five different sensors) and location information have been collected from 812 male and female drivers resulting in a databank measured in terabytes [4][6]. We have carried out some initial experiments using this CIAIR database.

The modalities used in the experiments are summarized in Table 1.

Table 1: Data Collection Specifications

Speech	Sampling: 16 kHz; 16-bit/sample; 12 channels
Video	MPEG-1; 29.97 frames per second; 3 channels
Driving Signals	Acceleration, Accelerator Pedal Pressure, Brake Pedal pressure, Steering Wheel Angle, Engine RPM, Vehicle Speed: Each at 16 bit/sample and 1.0 kHz.
Location	Differential GPS: one reading per second

4.1 Analyzing driving signals

In this section, we will present experiments related to verification of driving behavior of known drivers using driving signals available with the CIAIR database.

Driving signals have been used to verify whether a known driver is in normal driving condition using a subset of the CIAIR database. If the verification is not successful, this could indicate that there is a potential problem such as the driver may be fatigued or drunk. In this scenario, the impostor data for each driver should be gathered in fatigued or drunk driving conditions. However, CIAIR database does not contain such data. Thus, we assumed that the impostor data for each driver is given by the driving signals of the remaining drivers in the cross validation scheme. Details of the experiments are given in [4]

The experiments are carried out on two different subsets of the CIAR dataset, one having 28 drivers and the other having 314 drivers. Table 1 gives the unimodal success of each of the modalities on the dataset of 28 drivers, Table 2 gives the result of combining different modalities for the same subset. These tables show that gas and brake pedal signals have discriminative power. Classification performance of 69% has been reached in identifying drivers from these driving signals.
The performance results from the 314 driver subset of the CIAIR database show that, the performance decreases in the larger database, however, it is still observed that gas and brake pedal signals have discriminative power. Details are given in [4] These results are very encouraging in

that, the driving signals seem to carry important information to verify the driving behavior, and hence they may be used to detect fatigued or drunk driving conditions.

Table 2: Correct classification rates for 28 drivers (B: brake, A: gas, E: motor speed, S: vehicle speed, T: wheel angle, dX: derivative features).

Gauss mixture number	B/BdB	A/AdA	E/EdE	S/SdS	T/TdT
2	29.25 28.49	46.80 47.34	22.87 24.46	9.04 13.82	6.38 7.10
4	34.04 34.04	64.89 60.96	26.06 25.00	11.17 14.05	8.51 9.04
8	39.36 35.63	67.55 68.08	30.31 26.20	11.17 12.76	8.51 9.67

Table 3: Correct recognition rates for 28 drivers using combination of different modalities (B: brake, A: gas, E: motor speed, S: vehicle speed, T: wheel angle, dX: derivative features).

Gauss mixture number	B+A (1+2)	A+E (2+1)	B+E (2+1)	A+B+E (3+2+1)
8 (Bayesian)	64.36	63.29	44.68	68.61
8 (weighted)	69.14	68.08	46.80	68.61

4.2 Person Recognition Experiments

We have carried out open-set person identification experiments using audio and video from a subset of the CIAIR database. Details of the experiments are given in [4]

Our findings on the unimodal and multimodal experiments are presented in Table 4. The product rule, which assumes independence of modalities and combines the equally weighted modality scores, achieves 1.26% EER rate by improving the unimodal identification rates. On the other hand, the fusion of audio and face modalities with the RWS rule results in 1.04% EER rate by outperforming the product rule. These results indicate that combining audio and face modalities for open-set person identification considerably improves the overall performance.

Table 4: Open-set identification results using a 20 people subset of the CIAIR database, where ⊕ denotes the RWS rule and • stands for the product rule. A=audio, F=face

Modality	EER (%)
A	2.44
F	7.89
A•F	1.26
A⊕F	1.04

4.3 Fatigue Detection

The general-purpose facial expression analysis has been explored for decades. Numerous techniques have been proposed. A recent survey of existing works can be found in [7].

FACS representation of facial expression, is static and deterministic. Face expression develops over time and the detected facial features contain uncertainties. A dynamic and stochastic facial expressions representation framework is therefore needed.

Currently, a spatial and temporal model is being constructed to systematically represent and identify facial expressions like ‘inattention’, ‘yawning’, ‘falling asleep’.

5. Active Passive Restraint Systems

When a fatigued or influenced driver is detected, active safety systems should intervene and help the driver. In the limiting case of an over-fatigued or over-influenced driver, the vehicle control computer may stop the vehicle and automatically ask for assistance. In other cases, active safety features of the vehicle should help the driver. This may require temporary transition of driving control authority from the driver to the control computer. The simplest active safety features are of a warning nature and include alerting sound cues, or vibration of the steering wheel or driver seat to wake him/her up. Because of the potentially dangerous nature of these tasks, the active safety features are being developed in a simulator environment rather than using the test vehicles.

Realistic Vehicle Model Development

Since the active safety algorithms that are being considered are tested on simulators, realistic models of road vehicles need to be used. Along with the use of commercial road vehicle dynamics software, Simulink vehicle dynamics models at different levels of complexity were also prepared. The developed Simulink vehicle dynamics models range from single track to full vehicle models. Vehicle model has been illustrated in Figures 8.

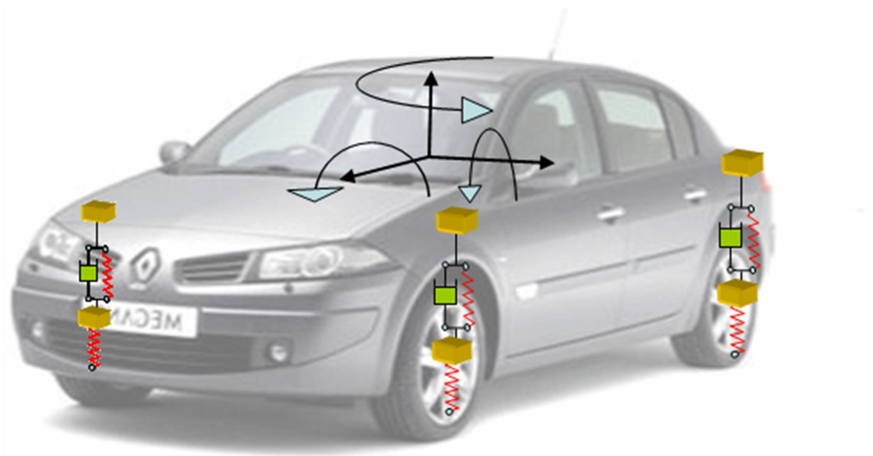


Figure 8. Vehicle model

Lane Following Assistance

In a preliminary study, a simple loss of attention control system was prepared in the simulator environment. This system analyzes the success of the driver in following the lane and intervenes if necessary with lane following assistance. A supervisory control architecture with two levels was used. The high level controller determines whether the driver needs lane keeping assistance or not. The high level controller is implemented using Stateflow and is rule based. If assistance is required, the low level automated lane keeping controller is switched on. This idea has been tested using a Simulink model of the vehicle with single track dynamics augmented with a Dugoff tire model, longitudinal dynamics and a static engine map model. The actual low level control action uses a PID controller.

Preview of the lane ahead is used along with the predicted trajectory of the vehicle to determine the successfulness of lane keeping. In the case of failure, an audio warning signal is first sent to the driver. If the driver still does not correct his lane following performance, the automated lane keeping controller takes over and keeps sending audio warning signals to the driver to correct his/her steering performance.

Other Active Safety Methods

Other active safety systems that have been developed and are being tested include yaw stability control, rollover avoidance and safe following of preceding traffic using adaptive cruise control and stop and go assistants. These control systems are being tested in two home built simulators. Another simulator to be used exclusively for active safety control testing in a driver attention monitoring and warning framework has been designed and is being constructed.

6. Summary and Conclusion

In this paper, we have presented the Drivesafe project whose goal is to create conditions for prudent driving on highways and roadways with the purposes of reducing accidents caused by driver behavior. We have discussed the different data collection efforts from multimodal sensors (such as cameras, microphones and other sensors) to build a unique databank on driver behavior.

We have given the results of preliminary studies on modeling driver behaviour using driving signals and multimodal person recognition technologies for secure personalized human-vehicle interaction.

Even though the performance of each individual modality could be increased in adverse conditions, such as using de-noising and echo cancellation for in-vehicle speech signal, the multimodal performance surpasses each uni-modal system once the best features or decisions from each modality are fused. Furthermore, multimodal person identification enables a fault tolerant design in case one of the sensors (e.g., one of the cameras or acoustic sensors) fails. Reliability measures can be assigned to each modality and/or to the output of each sensor so that only the features for the most reliable modalities and/or sensors can be considered in the fusion strategy. This includes the scenarios such as disregard speech modality when the background noise cannot be suppressed effectively, or disregard lip features if the driver is not looking straight into the camera, etc.

We also demonstrate, that the driving behavior signals can be used to verify current driving condition of an identified driver in a drive-safe scenario, where active/passive safety enforcement systems could be deployed if his/her behavior does not comply with pre-determined normal behavior.

In addition to constructing a spatial and temporal model to systematically represent and identify facial expressions like 'Inattention', 'yawning', 'falling asleep', potential future research directions include i) detection of best features for each modality, ii) optimum fusion strategies, iii) better behavioral modeling of drivers.

8. Acknowledgement

The authors would like to acknowledge Professor Fumitada Itakura of Meijo University and Professors Kazuya Takeda and Nobuo Kawaguchi of Nagoya University, Japan for providing the CIAIR database to be used in our case studies, State Planning Organization of Turkey and New Energy and Industrial Technology Development Organization (NEDO) of Japan for supporting the project.

KEY REFERENCES

[1] J. Campbell, "Speaker recognition: A tutorial," *Proceedings of the IEEE*, vol. 85, no. 9, pp. 1437.1462, September 1997.

[2] C. C. Chibelushi, F. Dearvi, and J. S. Mason, ``A review of speech-based bimodal recognition,'' IEEE Trans. on Multimedia, vol. 4, no. 1, pp. 23--37, 2002.

[3] P. Ekman, W. V. Friesen, and J. C. Hager, *Facial Action Coding System (FACS): Manual*. San Francisco, CA: CD Rom, 2002.

[4] E. Erzin, Y. Yemez, A. M. Tekalp, A. Erçil, H. Erdogan, and H. Abut "Multimodal Person Identification for Human Vehicle Interaction" *IEEE Multimedia,* Vol. 13, issue 2, pp. 18-31, 2006.

[5] K. Igarashi, C. Miyajima, K. Itou, K. Takeda, H. Abut and F. Itakura, “Biometric identification using driving behavior,” Proceedings of IEEE ICME 2004.

[6] N. Kawaguchi, S. Matsubara, I. Kishida, Y. Irie, H. Murao, Y. Yamaguchi, K. Takeda and F. Itakura, “Construction and Analysis of the Multi-layered In-car Spoken Dialogue Corpus,” Chapter 1 in *DSP in Vehicular and Mobile Systems,* Springer, New York, NY, 2005.

[7] M. Pantic and L. Rothkrantz, “Automatic analysis of facial expressions: The state of the art,” *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 22, no. 12, pp. 1424–1445, 2000.

[8] A. Ross, A. Jain, “Information Fusion in Biometrics,” *Pattern Recognition Letters*, vol. 24, pp. 2115-2125, 2003.

[9] W. Zhao, R. Chellappa, A. Rosenfeld and P.J. Phillips, “Face Recognition: A Literature Survey”, UMD CAR Technical Report CAR-TR-948, 2000.