Chapter 15
Computer Vision Systems for "Context-Aware" Active Vehicle Safety and Driver Assistance

Pinar Boyraz*, Xuebo Yang, and John H.L. Hansen

Abstract Recent developments of information technology and mobile lifestyle have forced drivers to multitask while they drive. The in-vehicle "infotainment" technology is already taking its place in the transformation of vehicles towards more intelligent and interactive devices rather than staying as mere transportation convenience. This transformation has several advantages such as easy route navigation, real-time traffic information, and staying connected with work or people while traveling. However, it has several drawbacks concerning the impact on driver cognitive load and attention sources. Therefore, it is crucial to take advantage of state-of-the-art in-vehicle technology to produce counter-measure systems that monitor the driver status and reduce driver workload adaptively depending on the context. In recognition and analysis of the driving context together with driver status monitoring, computer vision applications supply crucial information both in the vehicle (i.e., driver head and eye tracking) and out of the vehicle (i.e., lane, pedestrian, and vehicle detection and tracking, and road sign recognition). In this chapter, we provide a broad range of computer vision applications for CA-IVS from the literature and our previous studies, and we report our current research efforts.

---

* Pinar Boyraz was with the University of Texas at Dallas, CRSS UTDrive modeling group when this work was done. She has since joined Istanbul Technical University, Turkey.

P. Boyraz (✉)
Mechatronics Education and Research Center (MERC), Department of Mechanical Engineering, Istanbul Technical University, Inonu Cd. No.65, Gumussuyu, Istanbul 34437, Turkey

Center for Robust Speech Systems (CRSS), Department of Electrical Engineering
The University of Texas at Dallas, Richardson, TX 75080-3021, USA
e-mail: boyraz.pinar@googlemail.com

X. Yang • J.H.L. Hansen
Center for Robust Speech Systems (CRSS), Department of Electrical Engineering
The University of Texas at Dallas, Richardson, TX 75080-3021, USA
e-mail: john.hansen@utdallas.edu

J.H.L. Hansen et al. (eds.), *Digital Signal Processing for In-Vehicle Systems and Safety*,
DOI 10.1007/978-1-4419-9607-7_15,

Keywords Computer vision • Context aware • Lane tracking

15.1 Introduction

In their brief report, Fletcher et al. [1] provide an overall summary of promising computer vision systems applied in the vehicle. They determine areas where vision systems could be useful such as driver fatigue or inattention detection, pedestrian spotting, blind-spot checking, lane keeping, traffic sign recognition, and human factors aids. These applications are built based on several computer vision systems which are surveyed and/or presented here in this chapter as well. Building on what is already achieved in this area, we provide a systems engineering survey of computer vision systems for in-vehicle applications together with our previous and current findings. This study also presents a system utility analysis that ties all systems in a mechatronics integration approach, reducing complexity and cost of the final in-vehicle computer vision system, while maximizing the utility factor of the resultant design. In Sect. 15.2, applications are grouped into two main areas: driver status monitoring (inside the vehicle) and vehicle peripheral monitoring (outside the vehicle). These systems can be thought of as the *eyes of the cyber copilot* in the vehicle which (or who) is aware of driver's condition as well as the environment and the current situation (i.e., situation/context awareness). Next, in Sect. 15.3, all systems are analyzed from the perspective of utility in their *projected impact* on reducing the number of accidents or fatality rates. After determining the utility factors of systems, an example of mechatronics system integration for in-vehicle systems is presented. Finally, conclusions are drawn in Sect. 15.4 pointing to future research directions in this area.

15.2 Computer Vision Systems for In-vehicle Applications

In this section, we briefly survey different CV systems, reporting our progress in some areas with focus on the UTDrive research team. CV systems are seen as crucial components of future DAS and AVS systems; however, there is still a need for further development to achieve robust operation on board. Before providing details on each system, a list of requirements from onboard CV systems are presented here to emphasize the challenges in this area, some of which require hardware solutions and development of novel systems:

- Robust against illumination change
- Reliable in vibration and high accelerations
- Durable to low/high temperatures and weather conditions (especially cabling and mounting parts)
- Nonintrusive to the driver
- Compact/mobile
- Minimal power and computing source use

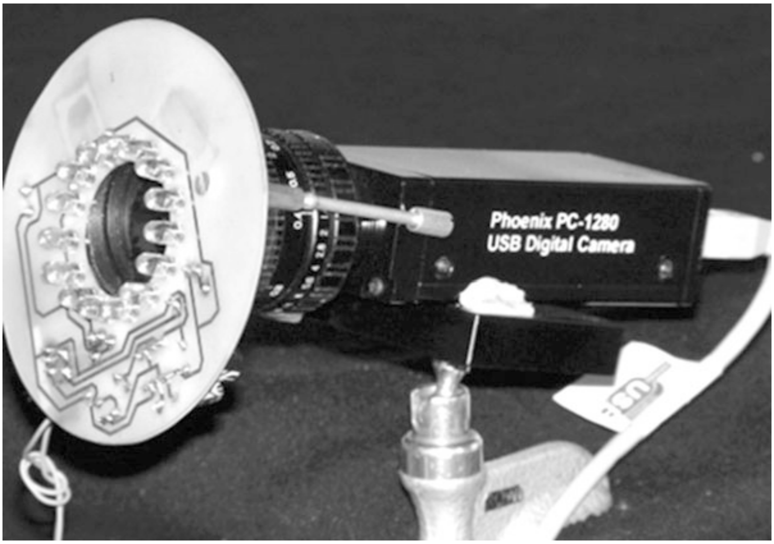

Fig. 15.1 NIR eye tracker designed as a part of driving monitoring system

Here, the CV systems are grouped under two main groups focusing on driver and on the environment covering all aspects of the driving context.

15.2.1 Eye and Head Tracking

Eye-tracking applications are originally motivated by research in human–computer interface development to create novel ways of interfacing [2] or helping people with motor disabilities [3]. In [4], an extensive survey of eye-tracking applications is given. For particular applications of eye tracking in driver monitoring systems, the systems can have a wide range from bright-pupil technique [5] utilizing co-centric near-infrared lights around the camera lens with active illumination to systems using an off-the-shelf webcam and visible light [6] and head-mounted systems [7]. There are also commercial eye-tracking applications already being used in studies utilizing eye-gaze information [8–10]. For head tracking, several applications using eye location, skin color, or motion in the image can be found [11,12]. A recent real-time system for monitoring driver vigilance is reported in [13]. In our previous study [14], an evolutionary computational approach was used to obtain an adaptive eye-tracking system to provide robustness in illumination changes. The system used bright-pupil technique which is based on retro-reflection property of the eye retina. The components of the system can be seen in Fig. 15.1, comprising a CMOS camera, co-centric ring of NIR LEDs to create bright-pupil effect, and optical absorption filters to block daylight.

Using the CV system in Fig. 15.1, eye tracking was performed to measure the pupil area as an indirect measurement of eyelid closure, eye gaze in x–y coordinate system, and head motion in 2-D image plane. The system can measure these three important indicators revealing drowsiness (i.e. PERCLOS [21] and [22]), attention level, as well as driver activity.

15.2.2 Affective Computing: Emotion Recognition

Emotion recognition can be multimodal using speech and video/image. The emotion recognition task is very difficult to achieve. It has been reported that even human coders are able to recognize the universal six archetypal emotions with accuracy of between 40% and 60%, especially when they are given the cues in single modality (i.e., only audio or only visual) [15], visible light [6], and head-mounted systems [7].

Although a tremendous amount of work exists in the face recognition area, emotion recognition remains a challenge since it has a temporal dimension as well and deals with nonrigid motion of the face. It is also a very young area which needs substantial work to reach the maturity of face recognition. However, there have been efforts to develop real-time and automatic units for emotion recognition using video modality. Anderson et al. [16] designed a fully automated multistage system for real-time recognition of facial expression. First, the faces are located using a spatial ratio template tracker algorithm, and optical flow of the face is subsequently determined using a real-time implementation of a robust gradient. The head motion is dealt with averaging and was canceled. The motion signatures from optical flow algorithm were classified using an SVM into non-expressive or six basic emotion types, as most of the work action units (AUs) were used. Shan et al. [17] investigated new subspace methods for reducing the features for facial expression analysis. Pantic et al. [18] especially emphasized the temporal characteristic of emotion sequences and had a detailed analysis of motion sequences using the profile face videos. However, they commented that this area needs to have a possible multi-camera system to deal with different viewing angles of the face and dynamic head motion cancelation. A survey of state-of-the-art automatic facial expression analysis can be found in Pantic et al. [19].

15.2.3 Vehicle Peripheral Monitoring

In this category, all road object detection and tracking systems can be included. Among them, the most promising systems are lane detection and tracking, road sign recognition, vehicle detection tracking, and finally pedestrian detection and tracking. Under the UTDrive research project, lane detection/tracking and road sign recognition systems are currently being developed with a context-aware framework.

15.2.4 Road Object Detection and Tracking

Video streams, whether processed online or off-line, contain rich information content regarding road scene. It is possible to detect and track vehicle, lane markings, and pedestrians and recognize road signs using a frontal camera and some additional sensors such as radar.

It is of crucial importance to be able to detect, recognize, and track road objects for effective collision avoidance or driver assistance system. In this chapter, we present our current progress in lane tracking and road sign recognition also reported in [22], adding a system utility analysis here.

15.2.5 Lane Detection and Tracking

There has been extensive work in developing lane tracking systems in the area of computer vision. These systems can be potentially utilized in driver assistance systems related to lane keeping and lane change. In [23], a comprehensive comparison of various lane-position detection and tracking techniques is presented. From that comparison, it is clearly seen that most lane tracking algorithms do not perform adequately so as to be employed in actual safety-related systems; however, there are encouraging advancements towards obtaining a robust lane tracker. A generic lane tracking algorithm has the following modules: a road model, feature extraction, post-processing (verification), and tracking. The road model can be implicitly incorporated as in [24] using features such as starting position, direction, and gray-level intensity. Model-based approaches are found to be more robust compared to feature-based methods. For example, in [25], a B-snake is used to represent the road. Tracking lanes in real traffic environment is an extremely difficult problem due to moving vehicles, unclear/degraded lane markings, and variation of lane marks, illumination changes, and weather conditions. In [26], a probabilistic framework with particle filtering was suggested to track the lane candidates selected from a group of lane hypotheses. A color-based scheme is used in [27]; shape and motion cues are employed to deal with moving vehicles in the traffic scene as well.

15.2.6 Road Sign Recognition

Methods used for automatic road sign recognition can be classified into three groups: color based, shape based, and others. The challenges in recognition of road signs from real traffic scenes using a camera in a moving vehicle has been listed as lighting condition, blurring effect, sign distortion, occlusion by other objects, and sensor limitations. In [28], a nonlinear correlation scheme using filter banks is proposed to tolerate in/out of plane distortion, illumination variance, background noise, and partial occlusions. However, the method has not been tested on different signs in a moving vehicle. Broggi et al. has addressed real-time road signs recognition in three steps: color segmentation, shape detection, and classification via neural networks; however, vehicle motion problem is not explicitly addressed. Jilmenez et al. used FFT signatures of the road sign shapes and SVM-based classifier. The algorithm is claimed to be robust in adverse conditions such as scaling, rotations, and projection deformations and occlusions.

15.3 System Utility Analysis and Mechatronics Integration

In recent years, whisper speech processing has attracted several researches. In this section, a system utility analysis is performed projecting the effect and cost of the surveyed CV systems onto 2007 FARS accident causation data [31, 32]. First, a query is run on FARS database to obtain the number of fatalities as the column and several driver-related factors on the rows. This table is rearranged into a more compact form and shown in Appendix 15.1. In this table, categories of causation are grouped under three major groups: driver impairment, driver errors, and in-vehicle devices. Redefining of these major groups in seven categories and matching them with appropriate CV systems that have the potential of preventing the accidents resulted in a new table shown in Appendix 15.2. The refined categories are: driver impairment, poor decision making, reckless driving, poor lateral control, poor longitudinal control, poor maneuvering, and in-vehicle devices. The distribution of the database is shown in Fig. 15.2. From this figure, it can be seen that only 34% of the fatalities are caused by driver-related factors; however, 66% of the data is unclassified and not reported clearly. Therefore, we may say that 34% is an underestimated figure. Nevertheless, the distribution within this 34% of the fatalities in terms of causation gives us important information about which types of driver errors should be prevented and where the drivers require the most assistance. From the distribution of the causation of accidents, we can clearly see that poor lateral and longitudinal control and maneuvering accounts for up to 65%. This figure can be reduced by proper DAS, warning, or active safety systems. Using the figures from the refined table in Appendix 15.2, a simple utility analysis is performed, and the results are shown in Table 15.1.

From the analysis results in Table 15.1, the most beneficial systems are determined to be lane tracking, optical flow, and traffic sign recognition. If an integrated system is used and integrated using the same sensor with modulation according to the imminent situation, the most beneficial system is traffic scene analysis. In the light of this justification, we report our recent efforts in designing a traffic scene analysis system with initial components being a lane tracker and traffic sign recognizer. The system

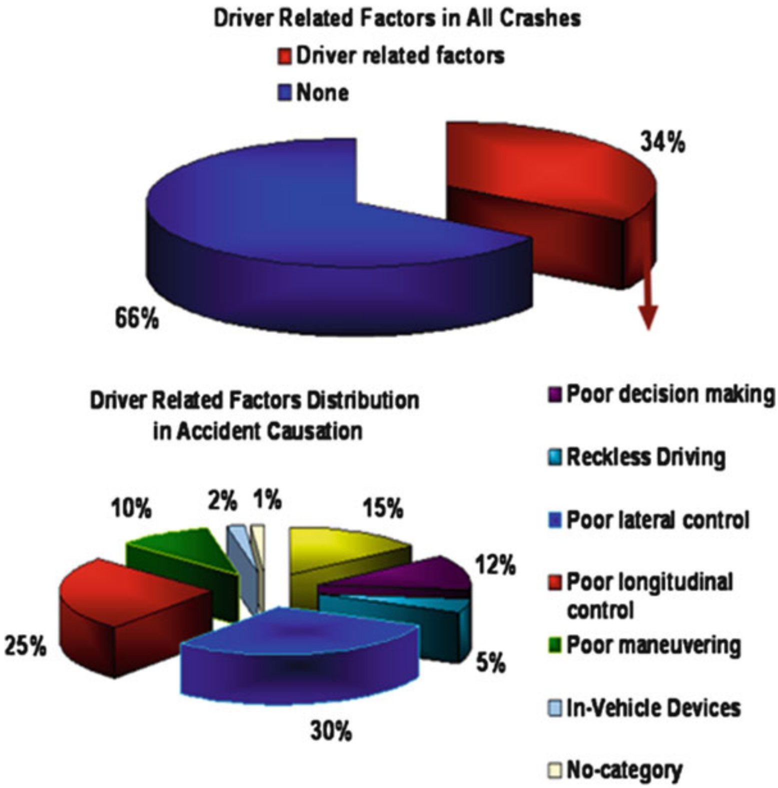

Fig. 15.2 Driver-related factors in crashes and its distribution

Table 15.1 Utility analysis results using projected prevention rate and unit cost of systems

CV system	Name	Projected prevention	%	Cost	Utility
Eye and head tracking	EHT	4,307	14.3	100	0.143
Emotion recognition	ER	1,683	5.6	100	0.056
Lane tracking	LT	10,304	34.3	100	**0.343**
Optical flow	OF	9,279	30.9	80	**0.386**
Lane change recognition	LCR	219	0.73	80	0.009
Road area recognition	RAR	66	0.22	50	0.004
Vehicle detection and tracking	VDT	1,585	5.28	100	0.053
Pedestrian detection and tracking	PDT	31	0.1	100	0.001
Traffic sign recognition	TSR	8,657	28.8	80	**0.36**
Integrated systems					
Traffic scene analysis	TSA	20,664	68.8	100	**0.688**
Driver warning system	DW	24,971	83.2	200	**0.416**

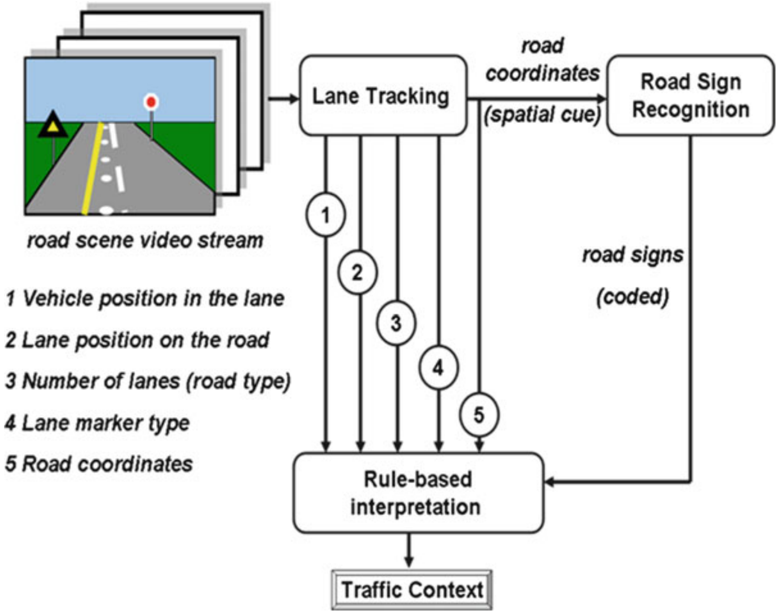

Fig. 15.3 General framework for TSA system. The details of the versatile lane tracker algorithm are given in Fig. 15.4

is presented in detail in [22]. The general framework is depicted in Fig. 15.3 with the aim to extract overall traffic context. The details of the lane tracking algorithm is given in Fig. 15.4. Sample outputs from road sign recognition module is shown in Fig. 15.5.

The fusion of information between different image processing modules can be realized using a rule-based expert system as a first step. Here, we present a set of rules combining the outputs of vision algorithms with the output options of warning, information message, and activation of safety features.

Case 1: If road sign is 0, standard deviation of lane position $<$ 10 pixels, standard deviation of vehicle speed $<$ 10 km/h, context: normal cruise.

Case 2: If road sign is 0, standard deviation of lane position $<$ 10 pixels, standard deviation of vehicle speed $>$ 10 km/h, context: stop-go traffic, likely congestion, and output: send information to traffic control center.

Case 3: If road sign is 1, vehicle speed $>$ 20 km/h, context: speed limit is approaching, output: warning.

Case 4: If road sign is 2, vehicle speed $>$ 20 km/h, context: stop sign is approaching and the driver did not reduce the vehicle speed yet, output: warning and activation of speed control and brake assist.

Case 5: If road sign is 3, vehicle speed $>$ 20 km/h, context: pedestrian sign is approaching, output: warning and activation *of brake assist*.

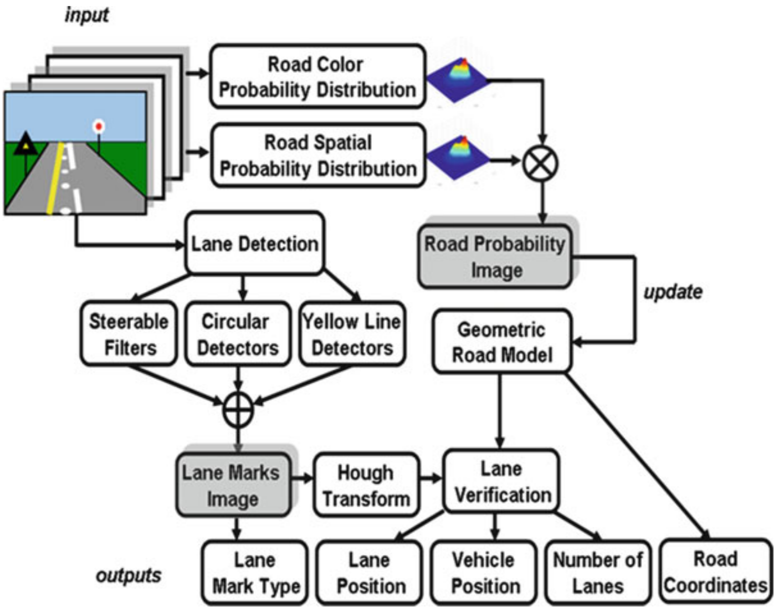

Fig. 15.4 *Versatile lane tracking algorithm.* Some example results from lane tracking and road sign recognition parts are shown in Fig. 15.5

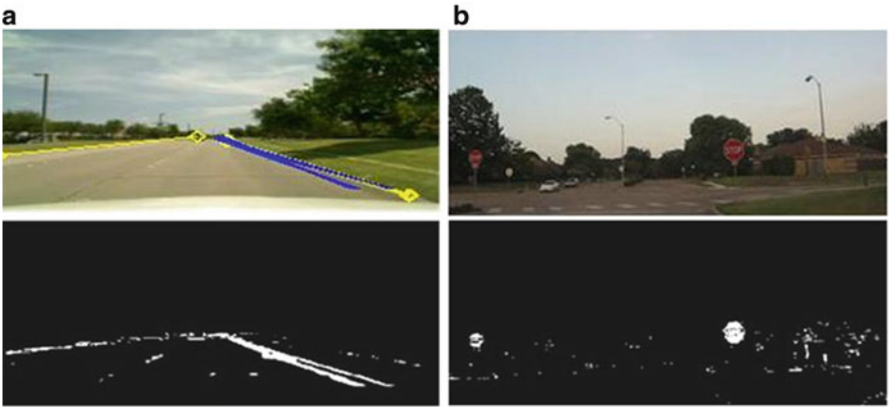

Fig. 15.5 (**a**) An example output of road area detection and lane tracker. (**b**) Color segmented stop signs after dilation in road sign recognition module

These cases represent only a subset of the rule-based schemes, and it is possible to use more advanced rule-based construction methods such as fuzzy logic (Figs. 15.4 and 15.5).

15.4 Conclusion

In this chapter, a brief survey of state-of-the-art computer vision systems for in-vehicle applications was presented. In a critical approach to gauge these systems with their benefits, a utility analysis is performed given that an integrated traffic scene analysis system would be the most optimal to work on. With the encouragement from the utility analysis, the recent efforts of UTDrive in combining different image/video processing algorithms with an integrated mechatronics approach using the same sensor were reported.

References

1. Fletcher L, Apostoloff N, Petersson L, Zelinsky A (2003) Vision in and out of vehicles. IEEE Intelligent Systems pp 12–17
2. (Online). www.fmcsa.dot.gov/documents/tb98-006.pdf
3. Morimoto CH, Mimica MRM (2005) Eye gaze tracking techniques for interactive applications. Comp Vis Im Und 98:4–24
4. Hutchinson TE, White KP Jr, Reichert KC, Frey LA (1989) Human–computer interaction using eye-gaze input. IEEE Trans Syst Man Cybern 19:1527–1533
5. Duchowski AT (2002) A breadth-first survey of eye tracking applications. Behav Res Meth Instrum Comput (BRMIC) 34(4):455–470
6. Zhu Z, Ji Q (2005) Robust real-time eye detection and tracking under variable lighting conditions and various face orientations. Comp Vis Im Und 98:124–154
7. Hansen DW, Pece AEC (2005) Eye tracking in the wild. Comp Vis Im Und 98:155–181
8. Sodhi M, Reimer B, Cohen, JL, Kaars R, Vastenburg E (2004) On-road driver eye movement tracking using head mounted devices. In: Proceedings of the Eye Tracking Research and Application Symposium (ETRA 2002), New Orleans. ACM Press, New York, pp 61–68
9. faceLab. http://www.seeingmachines.com/
10. Tobii Eye tracker. http://www.tobii.com
11. SR-Research. http://www.sr-research.com/
12. Xiao J, KanadeT, Cohn J (2002) Robust full motion recovery of head motion by dynamic templates and re-registration techniques. In: IEEE 5th international conference on automatic face and gesture recognition, Washington, DC, pp 156–162
13. Xu G, Sugimoto T (1998) Rits eye: a software-based system for realtime face detection and tracking using Pan-Tilt-Zoom controllable camera. In: Proceedings of fourteenth international conference on pattern recognition, vol 2. Brisbane, pp 1194–1197
14. Bergasa LM, Nuevo J, Sotelo MA, Barea R, Lopez ME (2006) Real-time system for monitoring driver vigilance. IEEE Trans ITS 6(1):63–72
15. Boyraz,P, Acar M, Kerr,D (2008) Multi-sensor driver drowsiness monitoring. IMechE PartD, J Automob Eng, 222(D11):2041–2063
16. De Silva LC, Pei CN, (2000) Bimodal emotion recognition. In: Proceedings of 4th IEEE international conference on automatic face and gesture recognition, Grenoble, 2000
17. Anderson K, McOwan PW (2006) A real time automated system for the recognition of human facial expressions. IEEE Trans Syst Man Cybern B Cybern 36(1):96–105
18. Shan C, Gong S, McOwan PW (2006) A comprehensive empirical study on linear subspace methods for facial expression analysis. In: 2006 conference on computer vision and pattern recognition workshop, New York City, New York. 17–22 June 2006, p 153, ISBN: 0-7695-2646-2

19. Pantic M, Patras I (2006) Dynamics of facial expression: recognition of facial actions and their temporal segments from face profile image sequences. IEEE Trans Syst Man Cybern B Cybern 36(2):433–449
20. Wierwille WW, Wreggit SS, Knipling RR (1994) Development of improved algorithms for on-line detection of driver drowsiness. In: International congress on transportation electronics (Dearborn), leading change. Warrendale, Society of Automotive Engineers
21. Knipling R, Rau P (2005) PERCLOS: a valid psycho-physiological measure of alertness as assessed by psychomotor vigilance. Technical Report of Federal Highway Administration, Office of Motor Crashes, USA FHWA-MCRT-98-006
22. Boyraz P, Yang X, Sathyanarayana A, Hansen JHL, Context-aware active vehicle safety and driver assistance, Enhanced Safety for Vehicles (ESV), 2009, 13–15 June, Stuttgart, Germany
23. McCall JC, Trivedi MM, Video-Based Lane Estimation and Tracking for Driver Assistance: Survey, System and Evaluation, IEEE Transactions on Intelligent Transportation Systems, vol. 7, no.1, pp. 20–37, March 2006
24. Kim YU, Oh, S-Y, Three-Feature Based Automatic Lane Detection Algorithm (TFALDA) for Autonomous Driving, IEEE Transactions on Intelligent Transportation Systems, vol. 4, no.4, pp. 219–225, Dec 2003
25. Wang Y, Teoh EK, Shen D, Lane detection and tracking using B-snake, Image and Vision Computing, vol 22, pp.269–280, 2004
26. Kim Z, Robust Lane Detection and Tracking in Challanging Scenarios, IEEE Transactions on Intelligent Transportation Systems, vol. 9, no.1, pp. 16–26, March 2008
27. Cheng HY, Jeng BS, Tseng PT, Fan, KC., ‘Lane Detection with Moving Vehicles in Traffic Scenes’, IEEE Transactions on Intelligent Transportation Systems, vol. 7, no.4, pp. 16–26, Dec 2006
28. Perez E, Javidi B, ‘Non-linear Distortion- Tolerant Filters for Detection of Road Signs in Background Noise’, IEEE Transactions on Vehicular Technology, vol. 51, no.3, pp. 567–576, May 2002
29. Broggi A, Cerri P, Medici P, Porta PP, Ghisio G, ‘Real Time Road Signs Recognition’, Proceedings of 2007 IEEE Intelligent Vehicles Symposium, Istanbul, Turkey, June 13–15, 2007
30. Jilmenez PG, Moreno G, Siegmann P, Arroyo SL, Bascon SM, ‘Traffic sign shape classification based on Support Vector Machines and the FFT of the signature blobs’, Proceedings of 2007 IEEE Intelligent Vehicles Symposium, Istanbul, Turkey, June 13–15, 2007
31. FARS Database: http://wwwfars. nhtsa.dot.gov/QueryTool/QuerySection/Report.aspx
32. Fatality Analysis Reporting System: http://www-fars.nhtsa.dot.gov/Main/index.aspx