Vision based Approaches for Driver Assistance Systems

IOAN BUCIU, ALEXANDRU GACSÁDI AND CRISTIAN GRAVA
Department of Electronics
University of Oradea
Universitatii Str., 1, Oradea, 410087
ROMANIA
{ibuciu, agacsadi, cgrava}@uoradea.ro http://webhost.uoradea.ro/ibuciu/

Abstract: - The number of cars on the roads increases with every passing year. Unfortunately, the number of accidents tends to be proportional with this increase. Nowadays, many vehicles are augmented with conventional driver assistance systems (DAS) such as: sensorial-informative (rear-view alarm, radar, lane departure warning system, etc.), actuation-corrective (anti braking system – ABS, emergency braking assistance – EBA, electronic stability control – ESC), and systemic (parking aids, distance control assist system, etc.). All aforementioned systems are already implemented and designed to improve the traffic safety and to, hopefully, reduce the number of car accidents (particularly fatal ones) mostly caused by human driver errors. Many efforts have been undertaken by joint collaborations between researchers and automotive industry to include Vision Systems on vehicle prototypes in order to enhance safety. For instance, to know the driver’s state, two cameras can be mounted on the dashboard to monitor the driver's gaze, triggering an alert if some drowsiness symptoms are detected. Driver gaze detection technology was thus developed. This paper presents an overview of the most representatives vision-based approaches developed to help the car driver in taking safe actions.

Key-Words: - vision systems, driver assistance systems, traffic monitoring, driver’s state monitoring.

1 Introduction

The sensorial-informative approaches are typically based on active sensors such as laser based sensor and millimeter-wave radar. However, the use of such sensors is limited. For example they cannot tackle the driver state. The driver’s drowsiness, inattention or fatigue is one of the major causes for car crashes that cannot be detected and predicted by using such sensors. Also, while active sensors can measure the distance between vehicle and obstacles to predict collisions, they cannot provide information about the type of obstacle. Another cause of accidents is provided by driving the car with higher speed than the allowed one prescribed by the speed limit road sign for difficult roads. A road sign recognition system would recognize the sign and warn the driver of exceeding the speed limit. This is another example where active sensors are unable to help. Therefore, some knowledge must be incorporated into DAS in order to enhance them with new capabilities. This can be accomplished by using Artificial Intelligence paradigms. These include fields such as: Image Processing, Machine Learning, Pattern Recognition and Human-Machine Interaction (interfaces). Many efforts have been undertaken by joint collaborations between researchers and automotive industry to include Vision Systems on vehicle prototypes in order to handle these challenges. For instance, to know the driver’s state, two cameras are mounted on the dashboard to monitor the driver's gaze, triggering an alert if some drowsiness symptoms are detected. Driver gaze detection technology was thus developed. Although promising progress has been accomplished in this direction (including the commercial systems), the topic is still in its infancy and plenty of issues remain to be solved or enhanced for designing more reliable and safety DAS. The paper aims at describing most advanced video-based approaches developed to be integrated into DAS.

2 Why Vision Systems for DAS

Vehicles can be equipped with two sorts of sensors: active or/and passive. Active sensors typically comprise laser based sensor and millimeter-wave radar. Passive sensors are represented by vision systems, where the information is acquired usually with the help of different types of video cameras. Both types of sensors have their advantages and limitations. The main shortcomings of active sensors refer to low spatial resolution and slow scanning speed. Another drawback is expressed by environmental pollution. Their use is limited by interference problems among sensors of the same types, which could become critical when large number of vehicles is running simultaneously on the same environment. Reflection waves occurring during data acquisition represent other active sensor limitation. One principal advantage of active sensors over the passive ones is determined by the fact that they can perform some measurements more accurate in a more direct way,
Table 1. Various sensor types

<table>
<thead>
<tr>
<th>Technology</th>
<th>Pros</th>
<th>Cons</th>
</tr>
</thead>
<tbody>
<tr>
<td>Laser</td>
<td>High angular resolution horizontally;</td>
<td>Optical components are sensitive against</td>
</tr>
<tr>
<td></td>
<td>Opening angle up to 360º;</td>
<td>pollution, rain or snow;</td>
</tr>
<tr>
<td></td>
<td>Fast data update rate;</td>
<td>No vertical resolution;</td>
</tr>
<tr>
<td></td>
<td>Accurate 3D depth data;</td>
<td></td>
</tr>
<tr>
<td></td>
<td>High distance resolution;</td>
<td></td>
</tr>
<tr>
<td>Radar</td>
<td>Relative low cost;</td>
<td>Problems in curves;</td>
</tr>
<tr>
<td></td>
<td>Resistant against pollution;</td>
<td>Interference problems;</td>
</tr>
<tr>
<td></td>
<td>High angular resolution horizontally and</td>
<td></td>
</tr>
<tr>
<td></td>
<td>vertically;</td>
<td></td>
</tr>
<tr>
<td>Image Processing</td>
<td>Possibility of an object-classification;</td>
<td>Difficulties in poor visibility;</td>
</tr>
<tr>
<td></td>
<td>Allows to detect forms, sizes;</td>
<td>Difficulties in case of strong variations of</td>
</tr>
<tr>
<td></td>
<td>Relative low cost for visible light imaging;</td>
<td>illumination</td>
</tr>
</tbody>
</table>

such as object movement. However, they are not able to identify obstacles, recognize road signs, and localize lane markers, etc., while vision systems are. Therefore, overall, the vision systems are definitely superior and obtain key advantages over the active sensors. Ultimately, a fusion strategy could be employed to benefit from both sensor types, exploiting one sensor type when the other sensor type either performs unsatisfactory or fails. Table 1 presents the sensors technology along with their pros and cons.

### 3 Vision-based Systems for DAS

Driver assistance system is a basic component of Intelligent Transportation Systems (ITS). The first ITS appeared more than 20 years ago, and, over time, the research community involved in this area known an increase interest in its effort to develop novel, more complex, trustable and accurate ITS. ITS technology has benefited from several disciplinary fields such as mechanical, electrical, electronic and, more recently, computer vision. Basically, a DAS can deal with one or several issues as follows:

1. Traffic monitoring;
2. Driver’s state monitoring;
3. Vehicle’s state monitoring;
4. Interactive driver – DAS communications;
5. Vehicle control;
6. Reasoning systems.

The first issue refers to tasks such as: preceding vehicle detection and tracking, static obstacles (buildings, trees, pylons, guard rails, etc.) detection and recognition, pedestrian and bicycle detection and tracking, traffic sign recognition, lane and zebra crossing detection, traffic light recognition. Being informed by a DAS about the traffic upcoming situations, a distracted or inattentive driver would pay increase attention at traffic conditions so that the potential collision incidents are hopefully minimized. The second component refers to tasks related to monitor the status of the driver: eyes detection, eye (gaze) tracking, head motion, detection of temporary lack of attention, stress detection, drowsiness and fatigue detection (eye blinking, yawning). The scope of this component is to predict if possible and to detect in real-time the driver hypo vigilance as this can cause road accidents. Vehicle’s state monitoring brings information about proper values of car’s oil, tire pressure, speed, cooling system, etc. which are crucial safety issues. An example of fourth component is given by the Global Positioning System (GPS) – based car navigation device, where the driver interactively communicate with the system (i.e. the driver can ask for some traffic information – addresses, shortest path to destination, way to avoid cluttered roads, etc., and the DAS provides requested information to the driver). Vehicle control is required when the DAS is augmented with capabilities of performing semi- or fully – autonomous maneuvers. Finally, a reasoning system should be able of performing tactical-level driving decisions.

Only the first 2 issues are presented here as they directly address the application of vision based techniques.

#### 3.1 Traffic monitoring

As noted above, traffic monitoring is related to:

1. Detection and tracking of preceding vehicles
2. Lane tracking
3. Pedestrian detection and tracking
4. Traffic sign recognition
5. Guard rail detection

#### 3.1.1 Detection and tracking of preceding vehicles

Broggi et al [1] proposed a simply method based on correlation between consecutive frames for detection and tracking of preceding vehicle. Vehicle is detected using edge detection algorithm where persistent edges around the candidate vehicle over consecutive frames are retained and considered as relevant candidates to detect and further to track the preceding vehicle. To speed up the
process, the analysis is performed over a window of interest, i.e. a sub-image. Some rules are defined to accept a detected front obstacle (represented through edges) as a preceding vehicle. Once the vehicle is detected, pixel correlation techniques are applied to successive frames for tracking the vehicle. However, the technique is applicable and effective only for slow moving of the preceding vehicle. Steux et al. [2] utilize a fusion scheme where a monocular color vision system is combined with a radar sensor to detect vehicles. Model-based approaches are used by Denasi and Quaglia [3], where the technique formulates an obstacle hypothesis by region segmentation algorithms and then validates it by matching an edge segmentation of these regions with a dynamic model. Fleischer et al. [4] also used a model-based technique to detect cars. Symmetry notion was investigated by several research teams such as Kuehnle [5] and Broggi et al. [6]. Batavia et al [7] proposed a technique to detect vehicles using an optical flow approach. However, optical flow method is sensitive to lighting conditions variations, and, besides, if the relative speed between the monitoring vehicle and the tracked one is low, the method fails. Fusing shadow, entropy, symmetry and texture-based approach is found in [8]. Stereo-vision system was involved in [9] where 3D features are extracted from image objects to predict the vehicle position. Moreover, by using a 3D feature extraction system an estimation of the distance between vehicles can be computed. However, this novel and recent technique suffers the same limitations as conventional edge-detection algorithms are used for feature extraction. Hilario et al. [10] came up with a geometric model defined by seven geometric parameters with optimum values found with the help of genetic algorithms. Generalizing, the aforementioned approaches can be also applied to a general case of obstacle detection. Finally some multi-sensory fusion strategies were employed to improve the accuracy of detection and to reduce false acceptance rate (defined as the percentage of identification in which false acceptance occurs, i.e. items that are wrongly classified as objects). Here the work described in [11] can be included.

3.1.2 Pedestrian detection and tracking

Among DAS issues, pedestrian detection and tracking is, perhaps, the most investigated component. Not surprisingly, since this issue faces the biggest challenges which are still open. The huge variety of pedestrian appearances makes the subject to be of great interest. People are dressed in various colors that sometimes blend with the background, making the segmentation process difficult. Moreover, the tracking step involves complex tracking algorithms that attempt to perform with a satisfactory performance. Different more or less affective techniques have been proposed to tackle this implying shape analysis and human gait. Active sensors are used such as laser scanners [12], passive sensors or fused sensors. Component based instead of holistic pedestrian representation is employed by Shashua et al. [13] to extract features further classified by a support vector machine (SVM). Gavrila and Munder [14] propose a multi-cue system where the detection phase involves a cascade of modules each utilizing complementary visual criteria to successively narrow down the image search space, balancing robustness and efficiency considerations. The method novelty expressed by the integration strategy of several modules as (sparse) stereo-based region of interest generation, shape-based detection, texture-based classification and (dense) stereo-based verification. The hierarchical structure of the system leads to performance close to the cutting edge techniques. Munder and Gavrila experimented in [15] several methods global and local feature extraction techniques for pedestrian detection, and found that local receptive fields combined with SVMs yields the highest accurate results.

3.1.3 Traffic sign recognition

While vehicle or pedestrian detection task can be preformed using only active sensors or fused sensorial architecture (active + passive sensors), the traffic sign detection and recognition is solely based on vision systems. Common approaches include color and / or shape based sign recognition techniques. For instance, Priese et al. [16] developed a system able to detect road sign based on their geometrical configuration and color. The work has been accomplished in collaboration with Daimler – Chrysler whose research groups put extensive research effort in this regard. Over years, they reported color based sign segmentation and parallel computing to tackle this task. However, the system is heavily dependent on color information, feature that can cause the system to perform poorly on images with weak or missing color information. Paclík et al [17] proposed a statistical expectation-maximization based approach where they focused on decision strategy employing Laplace kernel classifiers to develop a tree-classifier system. However, color information is again for segmentation process. A particular method has been investigated by Loy and Zelinsky [18] that uses local radial symmetry to detect radial shaped objects. The method is low computational and can be used in real-time applications. However, it can only be applied to concentric traffic sign detection since it relies on radial feature detection. The method is not applicable for other geometric shape of signs (such as triangular, square, etc).

3.1.4 Guardrail detection

Just a few papers are dedicated to guard rail detection. A representative work is describes by Broggi et al in [19],
where a method for detecting guard rails fusing radar and vision data in order to improve and speed-up vehicle detection algorithms is developed. The proposed strategy searches for uninterrupted oblique lines by involving a Sobel filter to detect edges followed by some rules (such as: horizontal edge components are kept whereas isolated points and vertical segments are erased), to make the algorithm more robust.

3.2 Driver’s state monitoring

Drowsy driving is a serious problem that leads to thousands of automobile crashes each year. In France, it causes about 30% of car crashes according to L. Arbus [20] and is responsible for one third of fatal crashes on French highways in accordance with ASFA (Assiociation of Highways French Societies) [21].

Researches on driver state monitoring has begun for about thirty years and are still very active. We can classify the driver state monitoring systems into three kinds of system:

- Systems that detect drowsiness through the vehicle behaviour.
- Systems that detect drowsiness through the driver physical behavior.
- Systems that detect drowsiness through the driver physiological behaviour.

The first systems which have been developed are the ones sensing the vehicle behaviour [22]. The main features studied are the steering wheel movements, the lateral position of the car on the road, the standard deviation of lateral position (SDLP) and the time to line crossing [23].

Research focused on systems sensing the driver behavior. The widespread technique to monitor the driver state is to use a video camera. Indeed, a lot of information can be extracted from the driver face to monitor fatigue such as the gaze, the frequency and the duration of eye blinking and yawning or the percentage of eyelid closure.

An interesting feature is the PERCLOS (PERcentage of eyelid CLOSure). This feature has been suggested by Wierwille [24]. It is the percentage of eyelid closure over the time. It allows detecting slow eye closures rather than blinks. PERCLOS has three drowsiness metrics. The most used metric is the P70 metric that is the proportion of time the eyes were closed at least seventy percent. The works of Dinges [25] and Knippling [26] show that PERCLOS is a good indicator of drowsiness and PERCLOS increases with the fatigue.

An interesting way of merging the different features (eye blinking, yawning, PERCLOS, gaze...) is used by Ji [27]. They use probabilistic networks that allow all features to contribute to the decision of the level of attention. Moreover extern factors (weather, hour of the day, etc.) can contribute to these networks and determine the level of attention.

Meanwhile the video features are not the best indicators of drowsiness. The best indicators of fatigue are the physiological indicators. In particular the electroencephalogram (also called EEG) is a good measure to monitor drowsiness. The EEG is the measure of the brain electric activity. The EEG is divided into five frequency bands (delta (0-4Hz), theta (4-8Hz), alpha (8-12Hz), sigma (12-16Hz), beta (16-32Hz)). The fatigue will be translated by an increase of the activity in the alpha and theta [28]. Other physiological indicators can also give information on the state of the driver such as the electro-cardiogram (ECG) or the electro-oculogram (EOG). The ECG gives the electric potential of the muscular activity of the heart.

3.2.1. Lane detection and tracking

Several strategies have been developed to tackle this issue. Some of them use windows of interest where only small parts of road regions (lane markings) where the lanes are most likely to be placed are analyzed [29].

Again edge detection procedures are employed combined with neural networks and filtering processing to detect the lanes. The strategy is characterized by fast processing and real-time performance. A histogram-based segmentation procedure is used in [30] to detect not only lanes but also objects on the road.

4 Open problems and challenging

Vision-based systems operating in visible spectrum used for traffic sign detection and obstacle and, particular, for pedestrian and vehicle detection suffer from various shortcomings. The main drawbacks are as follows:

1. Limited immunity to lighting conditions and shading. Most approaches used for these tasks are based on conventional edge detection methods which are not robust to change in illumination conditions. Both color and shape based approaches are prone to commit large errors when pixel intensity suddenly changes (obstacle highlighted by sunlight or cloud shadow). The presence of shadows produces artifacts in image representation thus negatively perturbing the detection system accuracy. More lighting condition immune systems should be found.

2. Limited and controlled test conditions. Most methods have been tested under regular weather conditions. While performing excellent in this case, their performance significantly decreases for bad and noisy weather situations, such as rainy or foggy situations. Noise resistant approaches ought to be investigated for more reliable systems performing under extreme weather conditions. Also, most systems have been tested on daytime with just a few work accomplished for night driving conditions.
3. Limited performance in the presence of occlusion. Occluded obstacles raise problems especially for model based or holistic based approaches.

4. Limited performance for complex road scenes. Satisfactory performances can be achieved with existing DAS for relatively simple road scene situations. The performances (detection rate and tracking accuracy for vehicle or pedestrian) rapidly decrease with an increase in road complexity. Detection and tracking tasks are much more difficult in cluttered and crowded scenes.

5. The challenge of pedestrian detection is many folds. At the detection level, one difficulty lies in distinguishing human bodies from other types of targets. This is a common problem for any sensor that is based on the reflection of acoustic, optical, or electromagnetic wave off a surface. Video image sensors suffer from the deterioration of performance in inclement weathers, poor lighting conditions, or special conditions. For example, shadows of objects in an image may cause confusion in identification or classification by computer vision techniques. At the tracking level, pedestrians may be stationary for no apparent reasons or they may move in any direction spontaneously. This makes the tracking problem particularly challenging. The image space variability of the pedestrian class is very large as they appear in various poses, clothing and various articulations of body parts. Pedestrians are found mostly in city traffic conditions where the background texture (from surrounding manmade structures, other vehicles poles and trees) forms a highly cluttered environment impossible to be discarded or ignored. The background clutter covers both shape (texture) and depth. If, for instance, a pedestrian would stand out in an open roadway using depth disparity cues (such as by using stereopsis), depth cues are unlikely to be useful for segmenting out pedestrians in city traffic due to the heavy disparity clutter. At tracking level, observing the pedestrian movement over time faces some drawbacks.

6. Although international standards exist for representing traffic signs, they actually may differ in their representation. There are several still open problems regarding driver’s status monitoring, as follows:

   1. Limited scenario applications.
   2. Still not negligible false alarm.
   3. Diagnosis performance needs improvements.
   4. High price and too many sensors causing problems during implementation phase.

Although, though many researches have been conducted about driver’s face monitoring with a camera, some limitations still exist:

   1. As for traffic monitoring, method present only limited immunity to lighting conditions, shading, night and day variations.
   2. Most of the works focused on fatigue detection and not really on inattention detection (for example, to be able to detect that the driver is no more looking at the road).
   3. A great challenge is to be able to be independent from head pose and head occlusion (face occlusion by a hand for example) and also to be independent from the driver.
   4. A great effort has to be done about predicting methods because with detection methods, it is often too late to make the driver reacting at time.
   5. Algorithms should take into account some models about the driver’s behavior in case of drowsiness or inattention situations. Especially, models about gaze movement in case of attention and inattention have to be studied.
   6. Real time processing is not always achieved.

5 Conclusion

The paper reviewed the most current approaches and techniques that may help and assist a car driver in improving the traffic safety from the vision point of view. More precisely, traffic monitoring oriented approaches and driver’s state supervision has been overviewed. Although impressive results were attained and some of those computer vision based approaches are already implemented on prototype or even real cars, many open problems still remain, especially when the approaches face unseen and difficult traffic scenarios. Major limitations of the current systems are also discussed.

ACKNOWLEDGMENT

This work was supported by a grant from the Romanian National University Research Council, PNCDI Program ID-668/2008.

References:


