Vision-based Driver Assistance Systems

Reinhard Klette
.enpeda.. (Environment Perception and Driver Assistance) Project
CITR, Auckland, New Zealand

5 February 2015

Abstract

Vision-based driver assistance systems are designed for, and implemented in modern vehicles, for improving safety and better comfort. This report reviews areas of research on vision-based driver assistance systems and provides an extensive bibliography for the discussed subjects.

1 Introduction

Vision-based driver-assistance systems (VB-DAS) still assume a driver in the *ego-vehicle* (i.e. that vehicle where the system operates in); autonomous driving can benefit from solutions designed for VB-DAS. Image or video data are recorded in the ego-vehicle and analysed for improved traffic safety or comfort. *Computer vision* is the general discipline for designing solutions for the understanding of image or video data [1].

1.1 What is Vision-based Driver Assistance ?

VB-DAS belong to the general class of *driver-assistance systems* (DAS). Besides cameras, DAS (often also called ADAS, for *advanced* DAS) use also further sensors such as GPS, IMU (= inertial moment unit), radar, sound, or laser range-finders [2].

See Fig. 1 for a sketch for multi-sensor data collection in a car. *Adaptive cruise control* (ACC) was pioneering those approaches in the early 1990*s*: longitudinal distance is measured by a radar unit (e.g. behind the front grille or under the bumper), and more recently also by employing a laser range-finder or stereo vision [3].

Historically, VB-DAS started with solutions for lane-departure or blind spot supervision; for example, see [4, 5] for related comments and references. It is now a common feature in modern vehicles; for example, see [6, 7].

VB-DAS combine one or multiple cameras, a processing unit with implemented applications, possibly interfaces to further sensors available in the vehicle, or to vehicle components related to vehicle control (e.g., if the system

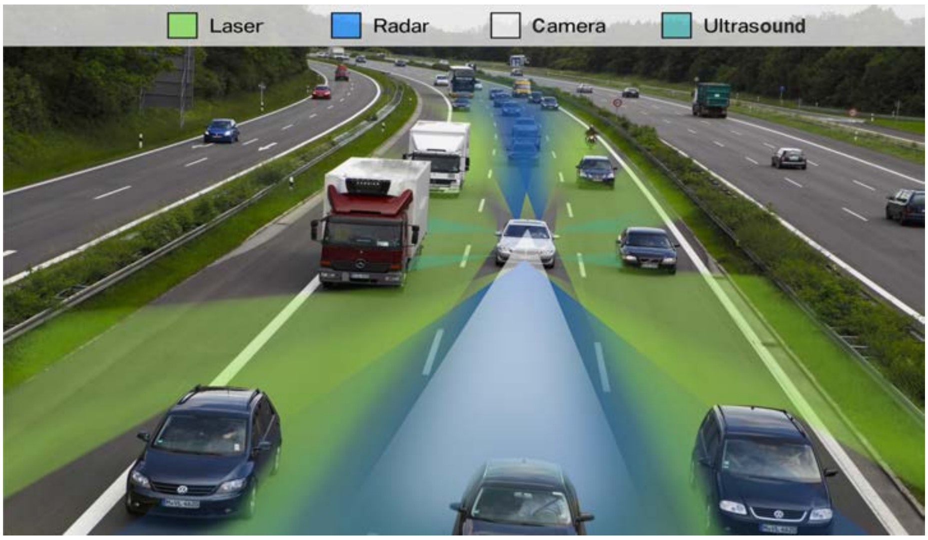

Figure 1: Graphical sketch for multi-sensor data recording. Courtesy of Clemens Dannheim

detects an obstacle on the left of the vehicle then steering to the left can be blocked) or to vehicle-driver communication (e.g. using the windshield as a head-up display, a visual signal may point to a potential risk).

This chapter discusses only tasks for computer vision related to driver assistance (*How to solve automatically visual perception tasks for driver assistance?*); it is not covering non-camera sensors or interfaces with other components in the ego-vehicle.

Computer vision often classifies approaches into *low-level vision* (e.g. image processing, stereo vision, or optic flow), *medium-level vision* (e.g. semantic segmentation of images, or object detection), and *high-level vision* (e.g. object tracking or the complex understanding of perceived scenes).

This classification does not correspond to the complexity of studied problems; for example, optic flow calculation (a low-level vision approach towards motion analysis) is in general more challenging than lane analysis for situations of well-marked lanes and reasonable lighting conditions, and often also more challenging than the detection and recognition of a particular traffic sign (e.g. stop sign or speed-limit sign).

1.2 Why Driver Assistance Systems?

In 2011, about 1.24 million people died worldwide due to traffic accidents [8], this is on average about 2.4 people *every minute*. Road injury ranked number 9 in 2011 for causes of death in the world, for example by far ahead of all war-related deaths.

VB-DAS, as DAS in general, are designed to reduce the number of traffic accidents (i.e. more safety for the driver, but also for all the passengers in the ego-vehicle, and also for any other participant in on-road traffic).

Improved comfort is the second reason for designing DAS. For example, stop-and-go driving at slow speed can be replaced by autonomous driving, or the unevenness of the road ahead can be detected and compensated by the car [6].

1.3 Proofs of Existence

The visual system of human beings provides a proof of existence that vision alone can deliver nearly all of the information required for moving around safely in the 3D world [9].

Visual odometry supported by our "build-in IMU" (i.e. accelerometer and gyroscope, with related sensors in the ears) defines human navigation. Thus, in principle, multiple cameras and an IMU are potentially sufficient for solving navigation tasks in the real world.

The Mars rovers "Curiosity" and "Opportunity" operate based on computer vision; "Opportunity" has done so since 2004.

1.4 Cameras and Frames

Cameras in the ego-vehicle record frames at different time slots t, typically at 25 Hz or higher frequency. We speak about *frame* t if these image data are recorded at time $t \cdot \triangle t$. Formally, frame t is denoted by $I(.,.,t)$ assuming a single camera for recording (i.e. a *monocular* case). The camera records grey-level or color values $I(x, y, t)$ at a pixel position (x, y) at time slot t.

A frame, recorded at time slot t, can also be a time-synchronized stereo-pair of images. In such a *binocular* case (i.e. like human stereo vision), we have two video streams; a frame is then composed of two images $L(.,.,t)$ and $R(.,.,t)$ for a left and a right channel, formally $I(.,.,t) = (L(.,.,t), R(.,.,t))$.

Outward-recorded (or *looking-out*) frames should have a large bit depth (e.g. at least 10 bits per pixel in each recorded channel), a high dynamic range (i.e. for being able to deal with "sudden" changes of light intensities between frames, or within one frame) and a high pixel resolution (e.g. significantly larger than just 640×480 VGA image resolution) for supporting accurate vision algorithms considering a wide horizontal angle (a wide vertical angle is less important).

Inward-recorded (or *looking-in*) frames for monitoring the driver, or the ego-vehicle occupants in general [10], might be of lower pixel resolution than for outward-recording.

Outward-recording should aim at covering as much as possible of the full 360° panorama around a vehicle. For doing so, multiple time-synchronized cameras are installed "around" the ego-vehicle, for example stereo cameras looking forward or backward. This extends binocular vision then to *multi-ocular* vision, and a recorded frame at time t is then composed of multiple images.

1.5 Performance Requirements and Evaluation

Ideally, VB-DAS have to operate for any occurring scenario, whether sunshine, rain in the night, driving in a tunnel or on serpentines, inner-city or highway traffic, and so forth. There are crash tests for evaluating the physical components of a vehicle; for evaluating solutions for VB-DAS it is necessary to run implemented applications on a large diversity of possibly occurring scenarios. Test data (i.e. a *video data benchmark*) can be recorded in the real world, or generated by computer-graphics programs (e.g. for simulating particular changes in processed frames).

The evaluation of solutions in relation to particular real-world scenarios has been discussed in [11]. Solutions can be characterized as being *accurate* or *robust*. *Accuracy* means correctness for a given scenario. *Robustness* means "sufficient" correctness for a set of scenarios which may also include cases of challenging scenarios. Ideally, robustness should address *any* possible scenario in the real world for a given task.

Used benchmarks should be of high diversity and complexity; used video data need to be evaluated for understanding their complexity. For example, *changes* in recorded video data can be characterized by using quantitative measures such as video descriptors [12] or data measures [13]. Sets of benchmark data should represent hours or days of driving in a wide diversity of possible scenarios.

Figure 2: Examples of benchmark data available for a comparative analysis of computer vision algorithms in a traffic context. *Left*: Image of a synthetic image sequence provided on EISATS (with accurate ground truth about distances and movements on this website). *Right*: Image of a real-world sequence provided on KITTI (with approximate ground truth about distances on this website)

Currently there are only very limited sets of data publicly available for comparative VB-DAS evaluations. Figure 2 illustrates two possible ways for generating benchmarks, one by using computer graphics for rendering sequences with accurately-known ground truth, as done for one data set on [14], and a second way by using high-end sensors (in the illustrated case for [15]; approximate depth ground truth is provided by the use of a laser range-finder).

1.6 Adaptive Solutions

We cannot expect to have all-time "winners" when comparatively evaluating computer vision solutions for defining VB-DAS applications. Vehicles operate in the real world, which is so diverse that not all of the possible event occurrences can be modelled in underlying constraints for a designed program.

Particular solutions perform differently for different scenarios, a winning program for one scenario may fail for another one. We can only evaluate how particular solutions perform for particular scenarios, possibly defining an optimization strategy for designing VB-DAS which is adaptive to the current scenario.

1.7 Premier Journals and Major Conferences

Research contributions in VB-DAS typically appear in premier journals such as *IEEE Transactions on Intelligent Traffic Systems*, *IEEE Intelligent Traffic Systems Magazine*, *IEEE Transactions on Vehicular Systems*, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, or *Computer Vision and Image Understanding*. Major conferences with contributions on VB-DAS are the annual *IEEE Intelligent Vehicles Symposium*, and the annual *IEEE Intelligent Transportation Systems Conference*.

2 Safety and Comfort Functionalities

Before discussing computer vision tasks, we briefly point to functionalities where recorded video data or graphical visualisations are simply used for enhancing the driver's visual perception of the environment, for safety or driver comfort.

2.1 Avoidance of Blind Spots

The *blind spot* is the total area around the ego-vehicle which cannot be seen by the driver. This is typically composed by an area behind the vehicle and two areas on the left and right of the vehicle. A simple VB-DAS solution is to show video data of those areas to the driver on a screen at times when of relevance.

2.2 Night Vision

VB-DAS may also support a driver's visual perception in the night or during otherwise limited viewing conditions (i.e. rain, snow, or fog), thus increasing the seeing distance and improving object recognitions. This is typically implemented by showing improved video data on a screen, but this can also be achieved by using a head-up display. Fog detection (for driver warning), see [16], is an example for distinguishing weather conditions.

The automotive industry designs active (i.e. use of a near-infrared light source built into the vehicle, which is invisible for the human eye but visible for a standard digital camera) or passive (i.e. no special illumination of the scene but capturing of thermal radiation) recording systems for providing enhanced images for the driver.

2.3 Virtual Windshield

The head-up display, or *virtual windshield*, is an efficient way for representing information to the driver without creating a need to change the head pose. Indoor-recording with face detection may be used for an accurate understanding of the driver's head pose.

The virtual windshield may be used, for example, for informing about speed, distance to destination, or navigation data. No computer vision needs to be involved for these cases.

The virtual windshield may also be used for informing about currently applying traffic signs (e.g. speed limit), for providing an enhanced view on lane borders in the night, for a flashing light indicating a view direction towards a potential hazard (e.g. a detected pedestrian in low-light conditions), or for labelling visible buildings (e.g. hotel chain).

For the mentioned examples, the shown information is mostly derived from

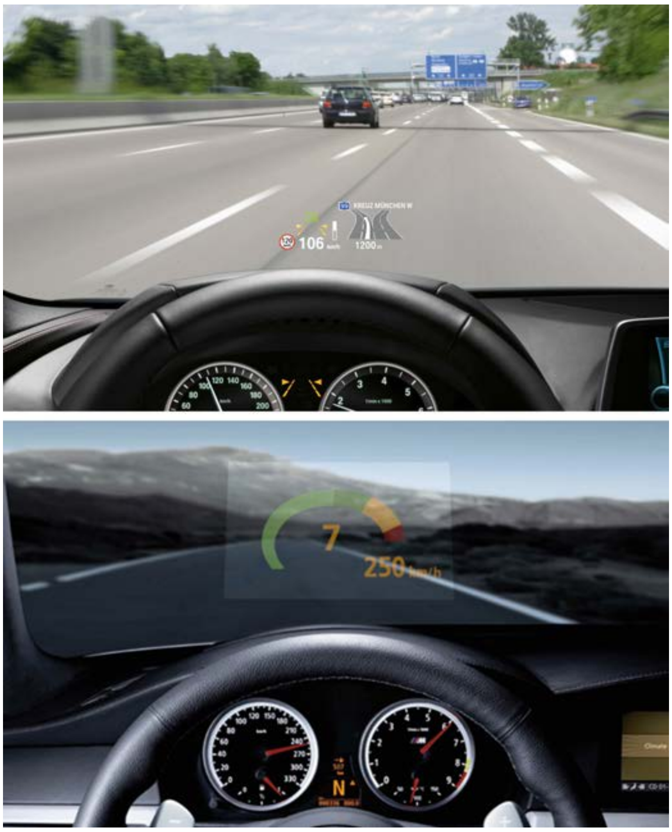

Figure 3: Two examples for a virtual windshield in a BMW. The "7" in the lower image identifies the current gear. Courtesy of Clemens Dannheim

particular VB-DAS applications (e.g. for traffic sign detection and recognition, visual lane-border analysis, or an advanced traffic hazard detection system; see related material later in this chapter).

3 Basic Environment Perception

The basic traffic environment consist of the ego-vehicle, other vehicles or pedestrians, traffic-relevant obstacles or signs, the *ground manifold* (i.e. the geometric ground-level surface; often it can be assumed that the ground manifold is approximately a ground plane, at least in a local neighbourhood to the ego-vehicle), the road, and the lanes.
Ego-motion describes the ego-vehicle's motion in the real world. Vision can help to control ego-motion according to the given obstacles or planned manoeuvres. For *basic navigation support*, only the existence of obstacles needs to be detected without understanding their type or movement, or the possible implications of those movements.

3.1 Distance Computation

Stereo vision is the dominant approach in computer vision for calculating distances. *Corresponding pixels* are here defined by projections of a surface point in the scene into images of multiple cameras. The applied vision system knows about the calibration data of those cameras and rectifies the recorded images into canonical stereo geometry such that 1-dimensional (1D) correspondence search can be constrained to identical image rows.

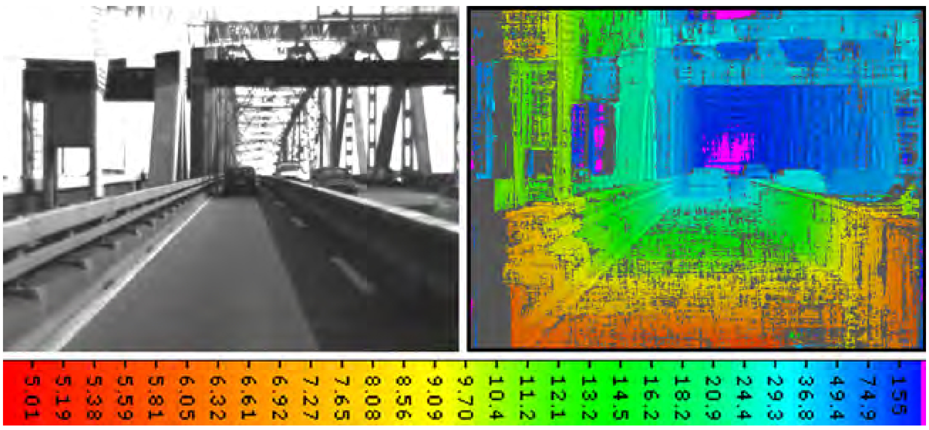

Figure 4: *Left*: One image of a stereo pair. *Right*: Visualization of a depth map using the colour key shown at the bottom for assigning distances in metres to particular colours. A grey pixel indicates low confidence for the calculated depth value at this pixel. Courtesy of Simon Hermann

Corresponding pixels define a *disparity*, which is mapped based on camera parameters into *distance* or *depth*. There are already very accurate solutions for stereo matching, but challenging input data (rain, snow, dust, sunglare, running wipers, and so forth) still may pose problems. See Fig. 4 for an example of a depth map. For example, stereo vision, combined with motion analysis (called *6D vision*, see [17]), provides basic information used in Daimler's "Intelligent Drive" system.

The third-eye technology [18, 13] provides a way for controlling the accuracy of an applied stereo matcher (i.e. of calculated disparity values). A measure based on *normalized cross-correlation* (NCC) is used for evaluating disparities frame by frame, thus identifying situations where a selected stereo matcher fails (and should be replaced by another matcher; see section on adaptive solutions above).

Combining stereo vision with distance data provided by laser range-finders is a promising future multi-modal approach towards distance calculations (recall that this chapter discusses vision sensors only). There are also ways to estimate distances in monocular video data [19].

3.2 Motion Computation

Dense motion analysis aims at calculating approximately-correct motion vectors for "basically" every pixel location $p = (x, y)$ in a frame taken at time slot t [1]; see Fig. 5 for an example. *Sparse motion analysis* is designed for having accurate motion vectors at a few selected pixel locations. Dense motion analysis is suitable for detecting short displacements (known as *optical flow*) [20], and sparse motion analysis can also be designed for detecting large displacements

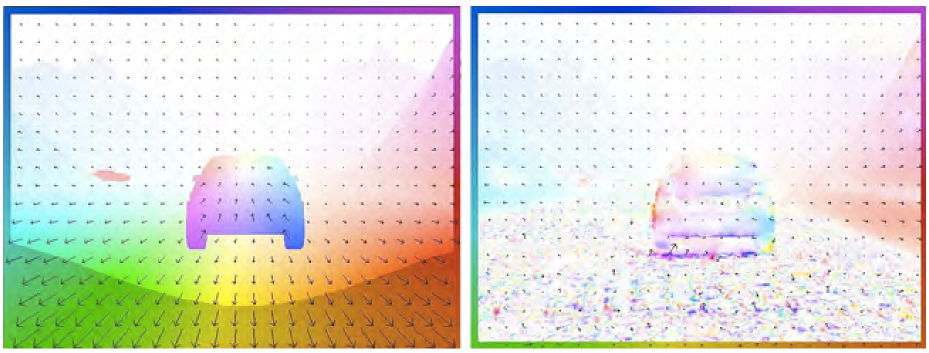

Figure 5: Visualization of optical flow using the colour key shown around the border of the image for assigning a direction to particular colours; the length of the flow vector is represented by saturation, where value 'White' (i.e. undefined saturation) corresponds to 'no motion'. *Left*: Ground truth for the image shown on the left of Figure 2. *Right*: Calculated optical flow using the Horn-Schunck algorithm published in 1981. Courtesy of Tobi Vaudrey

[21]. Motion analysis is a difficult 2D correspondence problem, and solutions might become easier by having recorded high-resolution images at a higher frame rate in future.

Moving objects in a traffic scene can be tracked by using repeated detections, or by following an object detected in a frame recorded at time t to a frame recorded at time $t+1$. A *Kalman filter* (e.g. linear, general, or unscented) can be used for building a model for the tracked motion as well as for involved noise [22]. A *particle filter* can also be used based on extracted weights for potential moves of a particle in particle space.

3.3 Ego-Motion

Object tracking is an important task for understanding the motion of the ego-vehicle, or of other dynamic objects in a traffic scene. *Ego-motion* needs to be calculated for understanding the movement of the sensors installed in the ego-vehicle. For example, an *inertial moment unit* (IMU) in the ego-vehicle provides a non-vision approach for ego-motion analysis.

Visual odometry uses recorded video data for calculating ego-motion; see Fig. 6. Possible approaches are characterized by *feature* tracking (a feature is a key point, i.e. a pixel, in one frame together with a descriptor characterizing

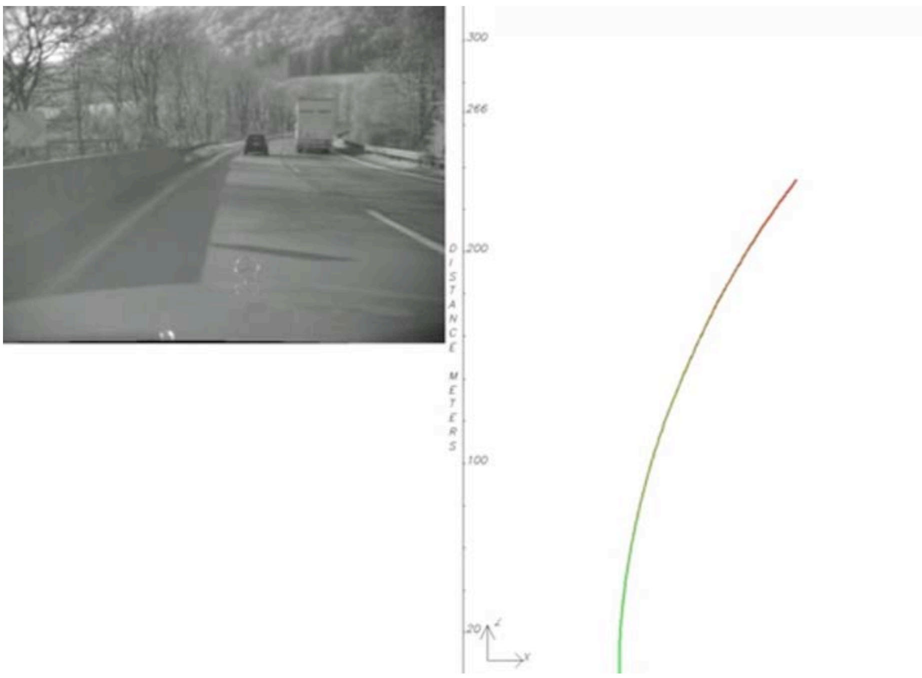

Figure 6: Calculated trajectory for the ego-vehicle of Sequence 3 in Data-set 1 of EISATS. Courtesy of Ali Al-Sarraf

image data around this key point; see [1]), *bundle adjustment* [23, 24] (i.e. the combined analysis of camera movement and of detected 3-dimensional points in the traffic scene), or by direct motion estimation, e.g. by simply applying an optical flow algorithm combined with non-visual sensor data such as GPS or of an inertial measurement unit (IMU) [25], or, more advanced, by applying 6D vision [17].

[24] defines *bundle adjustment* by refining the 3D model as well as detecting camera parameters. A set of n 3D points b_i is seen from m cameras (e.g. a camera at m different times while recording a video sequence). The cameras have parameters a_j. Let X_{ij} be the projection of the ith point on camera j. By bundle adjustment we minimize the reprojection error with respect to 3D points b_i and camera parameters a_j. This is a non-linear minimization problem; it can be solved by using iterative methods such as Levenberg-Marquardt.

3.4 Obstacle Detection

Monocular or stereo vision, often together with further sensors, provides input data for detecting vehicles, pedestrians, or further obstacles on the road [26].

For example, when applying stereo vision, detected points in the 3-dimensional scene need to be analysed for being just noise or actually obstacles on the road.

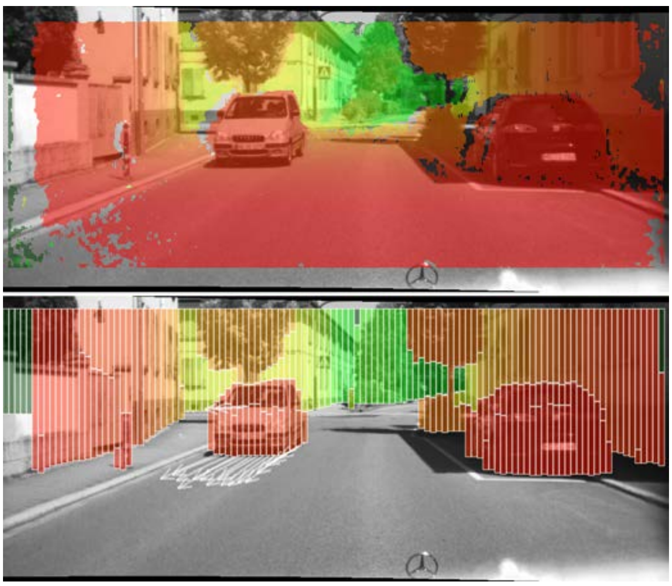

Figure 7: *Top*: Use of a color key (different to the one shown in Fig. 4) for showing depth data calculated by stereo matching. *Bottom*: Illustration of calculated stixels (based on the depth data illustrated above, forming an *occupancy grid*), groupings of stixels, and of estimated motion for such stixel groups. Courtesy of Uwe Franke

A detected local cluster of points at some height above the road define a *stixel* [27], which is a straight cuboid standing on an assumed ground plane and limited in height by the detected local cluster of points. *Regular stixels* are formed when assuming cuboids whose lower faces define a regular grid on the ground plane. See Fig. 7. Stixel classification can then aim at identifying basic object shapes like *car*, *bus*, *traffic sign*, or *construction cone*. Stereo vision also supports object detections on non-planar roads [28]. Generic object detection is studied in [29, 30], showing a good performance on [15].

Monocular object detection [31] has also been intensively studied for cases of monocular video recording (e.g. if attaching a mobile device to the windshield of a vehicle). [32] infer a vehicle (driving in front of the ego-vehicle) from the shadow underneath. [19] suggest a data-fusion approach using a boosted classifier (based on global Haar-like features) (in conjunction with corner and line features, and virtual-symmetry of tail lights of vehicles) to effectively detect vehicles, with a particular focus on also covering challenging lighting conditions. See Fig. 8. The book [33] discusses the detection of pedestrians in a traffic context; see also [34] and the database [35]. For a survey paper on pedestrian protection, see [36].

3.5 Detection and Tracking

There are static (i.e. fixed with respect to the Earth) or dynamic objects in a traffic scene which need to be detected, understood, and possibly further analysed. Typically, those objects are either the ego-vehicle itself, other on-road vehicles (e.g. also bicycles or children trolleys), or pedestrians.

Vehicle Tracking. Vehicle tracking is an important component of *collision avoidance systems*. By analysing trajectories of visible vehicles, in comparison to the trajectory of the ego-vehicle, it is possible to understand the danger of an imminent crash (e.g. to be used for triggering autonomous braking).

Tracking by repeated detection can use techniques as mentioned above in the section on vehicle detection. In general it is of benefit to use stereo vision results (i.e. disparity or depth values) in a vehicle tracking procedure [37], and not only monocular data.

Vehicle tracking is typically easier to perform than pedestrian detection and tracking; the shape and appearance of vehicles is easier to be modelled (e.g. by the appearance of lights, bumpers, horizontal line segments, density of detected corners, or visual symmetry; see [19]). Vehicle tracking is difficult due to occlusions, difficult lighting (e.g. light artefacts due to trees and intense sunshine), "ghost appearances" (e.g. reflected car headlamps on a wet road), and many more possible issues. Learning of ensembles of models has been proposed in [38], using data of [15] for training and testing. Supervised learning enhances the creation of a *discriminative part-based model* (DPM) from recorded video data [39, 40].

Pedestrian Tracking. Pedestrian detection, tracking, and understanding are in general still very challenging subjects. The task simplifies if only considering

Figure 8: Monocular vehicle detection under challenging lighting conditions. Detected vehicles are also labelled by monocular distance estimates. Courtesy of Mahdi Rezaei

pedestrians crossing the road, and not also pedestrians being close to the road (e.g. for understanding whether a pedestrian will step in the next moment on the road, or whether a child might possibly throw a toy on the road). A

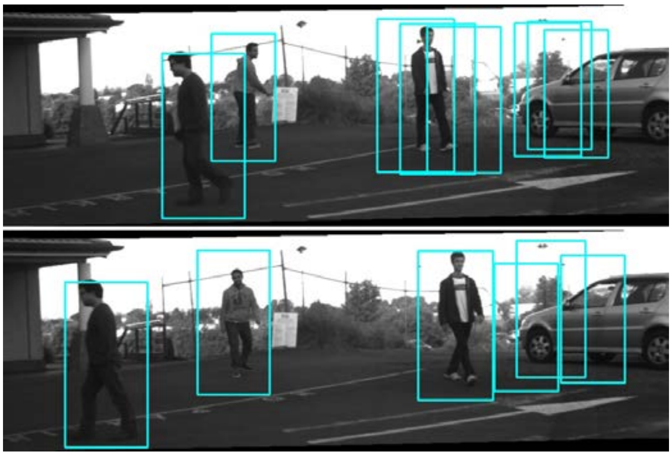

Figure 9: Two frames of a sequence of detected pedestrians using a RDF. There are a few false-positives (on the right in both frames), and also a few overlapping true-positive bounding boxes (in the upper frame for the person on the right). Further processing needs to eliminate false-positives, and to unify overlapping boxes. Courtesy of Junli Tao

straightforward approach for tracking is by repeated detection, possibly refined by taking previous detection results into account (up to Frame t) when analyzing Frame $t + 1$.

A standard procedure for detection is as follows: at first a *bounding box* (a window) is detected as the region of interest (RoI) which possibly contains a pedestrian. Apply a classifier for this bounding box for detecting a pedestrian. This classifier can be based on a *histogram of gradients* (HoG) for the bounding box [41]; after deriving *HoG descriptors*, the classifier uses those for deciding about the presence of a pedestrian. It is also possible to use such HoG descriptors within a *random decision forest* (RDF) [42] for performing the classification task.

For example, if the bounding box arrives at any leaf in the used forest which has a probability greater than 0.5 for the class "pedestrian", then the box may be classified this way (by this simple maximum-value rule). In case of overlapping bounding boxes, results may be merged into a single detection or box. See Figure 9.

Performance evaluation of pedestrian detection or tracking can be based on image data with manually identified ground truth; see, for example, the *Caltech Pedestrian Detection Benchmark* at `www.vision.caltech.edu/Image_Datasets/CaltechPedestrians/`. The *TUD Multiview Pedestrian* and the *CCV Pedestrian* databases can be used for body direction classification; they are available at `www.d2.mpi-inf.mpg.de/node/428` and `ccv.wordpress.fos.auckland.ac.nz/data/object-detection/` for free download.

3.6 Detection of Infrastructure Key Elements

The road, marked lanes, and traffic signs define the key elements of the traffic-related infrastructure of the environment.

Road Detection. Road detection is often considered as a pre-processing module prior to lane analysis [43], especially in cases of non-highway driving. [44, 45] discuss ways for modelling the visible road surface in front of the ego-vehicle.

The road might be identified by curbs, a particular surface texture, by a space between parked cars on both sides of the road, but also by very specific properties. See Fig. 10 for two extreme cases. In the case of roads in tunnels, walls and ground-manifold may have the same texture, and differ by there surface gradients. In case of an unpaved road in a desert, the texture of the road surface may continue on the left or right, and only traces of previous driving may indicate the actual location of the road.

Figure 10: *Left*: Roads in tunnels (below the historic centre of Guanajuato). *Right*: Unpaved road near Salta in Argentina. Courtesy of authors of [46]

Lane Analysis. In a general sense, a *lane* is defined by sufficient width for driving a road vehicle; it is the space between a left and a right *lane border*. Many different mathematical models have been used for defining lanes (e.g. analytically defined curves or sequences of individual border points following some kind of systematic pattern). In the simplest case, straight segments are used for describing zero-curvature lane borders, and second order curves or clothoids for non-zero-curvature lane borders.

There is already a vide variety of solutions available for lane analysis; see [46, 47, 48, 49, 50].

Lane detection is basically "solved" for scenarios during driving where lane markings, lane geometry, and visibility conditions are reasonable, but there is still a need for studying lane-border detectors or trackers for challenging scenarios (e.g. underground road intersections, unpaved roads, or very wide road intersections without any lane marking).

There is also not yet any satisfying automatic evaluation available for quan-

Figure 11: Three examples for data provided by [52] where a used lane detector follows its strategy for detecting lane borders due to temporal inference, but where one given frame alone would be insufficient for a judgement whether the detected border is correct or not. The three images show detected lane borders composed of sequences of individual points. Courtesy of Bok-Suk Shin

tifying the performance of a lane detector. For example, we could claim that "*lane borders are correctly detected if they are within an error of at most 5 cm to the true lane border*". What exactly is the "true lane border"? How to measure for cases as illustrated in Fig. 11? [15] offers a few manually-labelled frames for evaluating lane detection. Synthetic data for evaluating lane border detectors are available on [51].

The detection of lane borders is sometimes even a challenge for human vision. Lane borders can often not be identified in an individual frame; see Fig. 11. Additional knowledge such as the width of the car or the previous trajectory of the car can be used for estimating the continuation of lanes.

[52] proposes a semi-automatic technique for generating ground truth for lane detection. They use *time slices*, being defined by taking a specified single row with detected lane locations in subsequent frames, and fit splines to the resulting sequences of individual points in such time slices. By specifying different rows, different time slices are created. The proposed approach works reasonably well on clearly-marked roads. The involved interaction comes with the risk of human error and limited usability.

Traffic Signs. *Road signs* are traffic signs (stop sign, speed sign, etc.) or any form of written (directions, weather conditions, closure times of a lane etc.) or graphically expressed information (pedestrian crossing, speed bump, icons, etc.) on or near the road which are of relevance for driving a vehicle on this road. Classes of road signs can define one particular module of a complex computer-vision system for ego-vehicle control. For a survey on traffic sign detection, see [53].

A standard approach [54] can be briefly sketched as follows: possibly pre-process an image (e.g. by mapping a colour image into HSV colour space [1]), detect geometric shapes (circles or polygons) which are potential candidates for a traffic sign (possibly using colour as a guide as well), extract features, and compare those with features of a data base of traffic signs.

Solutions can be classified in general by focusing either more on the use

Figure 12: *Left*: Detected relevant features in an input image. *Middle*: Detected sign due to voting by SIFT features which passed the potential location filter. *Right*: Diversity of the appearance of the P30 sign in New Zealand. Courtesy of Feixiang Ren

of colour, or more on the use of shape for the initial detection. For example, circles can be detected by using a Hough transform [1] or a radial-symmetry approach [55]. Recorded images are subdivided into regions of interest (i.e. left or right of the road, or on top of the road) for having size-priors for traffic signs in those regions. See Fig. 12, left, for a case when detecting image features uniformly, all over the image, in the middle when restricting the search to regions of interest, and on the right for illustrating the diversity of traffic signs. Traffic sign categorization is a main subject in [54].

The authors of [56] suggest an evaluation methodology for traffic sign recognition by specifying measures for comparing ground truth with detected signs. Of course, before applying this methodology the ground truth needs to be available, and so far it is provided manually. GPS and e-maps allow us to compare locations of detected traffic signs with mapped locations of signs.

Free Space Detection. *Free space* is the area where the ego-vehicle may evolve safely. [57] is an example of early work for detecting free space based on color analysis.

More recent solution in VB-DAS use stereo vision for calculating occupancy grids [44, 45, 58]. See Fig. 7, bottom, for a stixel-illustration of an occupancy grid.

4 VB-DAS Examples

A traffic scene is composed of the road (with lane markings, pedestrian crossings, speed bumps, cavities, and so forth), *road furniture* (e.g. traffic signs, handrails, or construction site blocks), various types of obstacles (driving vehicles, pedestrians, rocks, parked vehicles, and so forth), and also by traffic-related buildings, tunnels, bridges, and so forth.

A complex scene analysis needs to understand the current traffic-related components, their motion, and their possible near-future impacts on the ego-vehicle or even the possible impacts on other traffic participants. Before discussing traffic scene-analysis tasks for cameras which point outwards from the ego-vehicle, we consider at first a few tasks when pointing a camera towards the driver (for understanding awareness).

4.1 Driver Monitoring

Cameras are not the only way for driver monitoring. For example, see [59] for a tactile solution using an embedded sensor in the steering wheel. Cameras are not only useful for understanding the state of the driver (e.g. drowsiness detection) but in particular also appropriate for analysing the viewing direction.

Face and eye detection [60], or head-pose analysis [61] are basic tasks in this area. The viewing direction can be estimated on head pose analysis; eye gaze analysis [62] is an alternative way for viewing direction estimation which also covers eye state analysis (i.e. percentage estimate for being open or closed). Challenging lighting conditions still define unsatisfactorily-solved scenarios; for example, see [63] for such scenarios. Foot gesture or visual hand motion patterns are further possible indicators for driver monitoring [64, 65].

Driver awareness can be defined by relating driver monitoring results to environment analysis for the given traffic scenario. The driver not only needs to pay attention to driving; eye gaze or head pose [19] should also correspond (for some time) to those outside regions where safety-related events occur. Here,

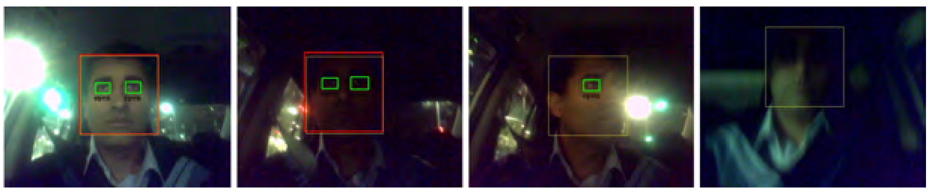

Figure 13: Face detection, eye detection, and face tracking results under challenging lighting conditions. Courtesy of Mahdi Rezaei

head pose or face detection is then typically followed by eye detection and an analysis of the state of the eyes or eye-gaze detection. See Fig. 13.

Figure 14: Detected eyes in a (supposed to be) driver's face based on defining a region of interest within a detected face for expected eye locations. *Right*: Examples of local Haar wavelets. Courtesy of Mahdi Rezaei

In the applied face or eye detection technique [60], a search window scans through the current input image comparing intensities with *local Haar wavelets*; see Fig. 14, right, for examples of such wavelets. [63] also introduces *global Haar wavelets*, motivated by challenging lighting conditions (as illustrated in Fig. 13). Figure 14 also illustrates the use of a head model for identifying eye regions.

4.2 Speed Adaptation

Now we mention a first task for outward-recording cameras. *Intelligent speed adaptation* (ISA) can be based on knowing the currently-applying speed limit, road conditions, the distance to the vehicle in front, and further traffic-related events (e.g. children playing close to the road often require a speed reduction).

Basic information for ISA is available in digital maps via GPS. Vision technologies are adequate for collecting on-site information. The detection and interpretation of speed-signs [66] or road-markings [67] are examples of traffic-sign analysis, and the evaluation of road conditions is part of road environment analysis (see related paragraphs below).

4.3 Queuing

An *automated queue assistant* (AQuA) applies in congested traffic situations on highways; see Volvo's program for AQuA design in [68]. For example, it is of interest for a truck convoy to maintain constant speed and distances between trucks. But it is also of importance for any congested traffic situation.

An AQuA application should ideally combine longitudinal distance control (to the preceding vehicle) for adjusting speed and lateral control for steering supervision (i.e. distance to side vehicles). Driver monitoring (see paragraph above) is also of significance for understanding driver awareness (drowsiness or inattentiveness). Lane detection and analysis (e.g. of lane curvature; see paragraph above) is of importance for proper positioning control of the vehicle.

For example, a truck convoy may be grouped for automated driving into a *platoon*, with the goal to reduce inter-truck distances for increasing the capacity of roads.

4.4 Parking

Automated parking requires in general a wider field of view than blind spot surveillance, and a full 360° recording supports both applications [69, 70].

Autonomous parking systems use typically multiple sensors such as ultrasonic or close-range radars, or laser scanners (implemented in the front and rear bumpers of the car), and vision sensors. The system needs to detect a possible parking space, and needs to guide the vehicle autonomously into this space. Note that bumper-implemented sensors do have a limited field of view (i.e. close to the ground).

Automated parking is currently extended to solutions for parking autonomously in a parking house with the drop-off and collection point at the entrance of the parking house. Vision sensors play an essential role in such an application.

4.5 Blind Spot Supervision

At the beginning of the article we already mentioned blind spot visualization. More advanced applications analyse the video data recorded for blind spots, and communicate only the necessary information (e.g., there is another vehicle on the left) to the driver [71]. For example, the *Blind Spot Information System* (BLIS), introduced by Volvo in 2005, produced just an alert if another vehicle was detected left or right of the ego-vehicle (by analysing recorded video); see [72].

4.6 Lane Departure Warning

Lane analysis (as discussed above) aims in particular at providing information about lane changes, if a driver is interested in this type of support (e.g. truck or long distance bus drivers).

4.7 Wrong Lane Detection

Wrong-roadway related accidents lead only to a relatively small number of accidents, but with a high risk of being a heads-on crash. About 33.6 percent of the world population drives on the left-hand side.

Lane-positioning algorithms based on e-maps and GPS [73, 74, 75, 76] are related to wrong-lane detection; this *map matching* approach is not yet accurate due to existing variance in (standard) GPS data. A lane-positing module together with a lane-detection and tracking module allow us to design a wrong-lane detection system.

The location, provided by GPS, can be matched with an available e-map, for instance, an *openstreet map*. The number of lanes and lane direction(s) at the current location needs to be read from this map. Apart from using GPS

Figure 15: Three detected lanes in a country driving on the left-hand side. Matching with an available e-map simplifies the decision about the lane the ego-vehicle is currently driving in. Courtesy of Junli Tao

and an e-map, further sensors such as onboard odometer or gyroscope can be used to refine the accuracy of ego-vehicle positioning on a road [76].

A multi-lane-detection result (e.g. as shown in Fig. 15) is mapped onto a current lane configuration, thus supporting (besides a detection of the central marking in the middle of the road) the decision in which lane the ego-vehicle is currently driving in. Methods as described in [77, 78, 79] address multi-lane detection. Lane confidence measures can be used to weight detected lanes for producing stable detections [80].

A first assistance system for the detection of driving on the wrong side of the road by reading no-entry signs of motorways is reported in [81], and [82] analyzes motion patterns in highway traffic for understanding wrong-lane driving. [80] propose a system for wrong-lane driving detection by combining multi-lane detection results with e-map information.

4.8 Intelligent Headlamp Control

Adaptive LED headlamps adjust the individual light beams according to the visible traffic scene [83]. Each light beam (of each addressable LED emitter) may vary between low-aimed low beam and high-aimed high beam. The beam is adjusted for maximizing the seeing range for the driver in the ego-vehicle, but with the constraint to avoid dazzling drivers in other vehicles or pedestrians.

Cameras in the ego-vehicle are used for vehicle and pedestrian detection (for these topics see later in this chapter) for changing permanently a *glare-free high beam* pattern of the headlamps.

5 Future Developments

Vision technologies, in combination with further technological developments, offer various new opportunities to improve safety and comfort; we mention a few.

5.1 Driver-Environment Understanding

Sections above discussed the understanding of the driver (i.e. inward-recording for analysing awareness, eye gaze, and so forth), and also various modules of VB-DAS for outward-recording. The obvious next step is to correlate driver understanding with traffic scene understanding, for example for warning about a detected issue for which it appears that the driver did not yet pay attention [19, 68]. For example, the virtual windshield (i.e. a head-up display) appears to be a good implementation for such a warning.

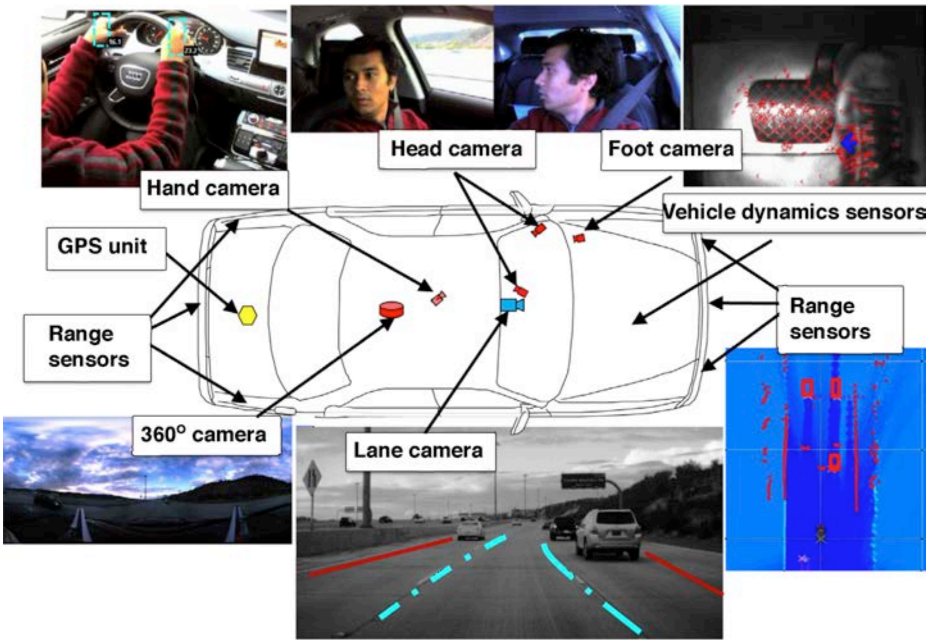

Figure 16: Holistic scene understanding and driver monitoring using multiple sensors for inward and outward recording. Driver awareness can be modelled based on a combined analysis of those data. Courtesy of Mohan Trivedi

Approaches for VB-DAS which combine looking-in and looking-out techniques provide significant benefits for realizing active safety functionalities; see [84, 85, 86, 87, 88] and Fig. 16.

5.2 Inter-Car Communication

Inter-car communication (a non-vision technology) supports the better understanding of large-scale road environments, not apprehendable from the perspective of a single car. The approach is also known as *car-to-car* (C2C). It is expected that these communication networks will contain vehicles in a local neighbourhood but also roadside sources as nodes (C2I = *car-to-infrastructure*).

Communicated information will include data collected via VB-DAS, and also data collected via stationary cameras along the road. Inter-car communication will be part of expected *intelligent transport systems* (ITS), defining the general context of future transportation.

5.3 Autonomous Driving

Autonomous driving is often identified as being the ultimate goal when designing VB-DAS [89, 90]. There are already convincing demonstrations that autonomous driving is possible today using, e.g., dominantly stereo vision [91] integrated into a car, or using dominantly a laser range-finder [92] mounted on top of a car (together with very accurate e-map and GPS data). In both cases, environment and weather conditions have to satisfy particular constraints for guaranteeing that the systems work accurately.

5.4 Road Environment

When analyzing the road environment, we are typically interested in detecting and interpreting road furniture, pedestrian crossings, curbs, speed bumps, or large-scale objects such as an entrance into a tunnel, or a bridge. Accurate e-maps and GPS provide information about the expected environment; cameras and computer vision can be used for detecting unexpected changes in such an environment.

Ground-level recording and 3D reconstruction (i.e. only using cameras in vehicles, not also aerial recording) is a subject, for example, in [93, 94, 95], all using multiple cameras while recording road sides in a *single run* when moving the cameras in the ego-vehicle basically parallel to the road borders into one direction only (i.e. without any significant variations in the path).

3D road-side visualization or 3D environment modelling are applications which still lie beyond the current interest in VB-DAS. A 3D reconstruction from a moving platform [93], possibly in combination with 3D reconstructions from a flying platform such as a multi-copter, can be used for an even more accurate environment model compared to today's planar e-maps.

3D surface data, reconstructed at time t, need to be mapped into a uniform world-coordinate system. Generating accurate surface models by mapping reconstructed 3D data requires a very high accuracy of ego-motion analysis. This accuracy is currently not yet available. This is in particular apparent when trying to unify results from different runs through the same street [96]. See Fig. 17 for 3D results from a single run.

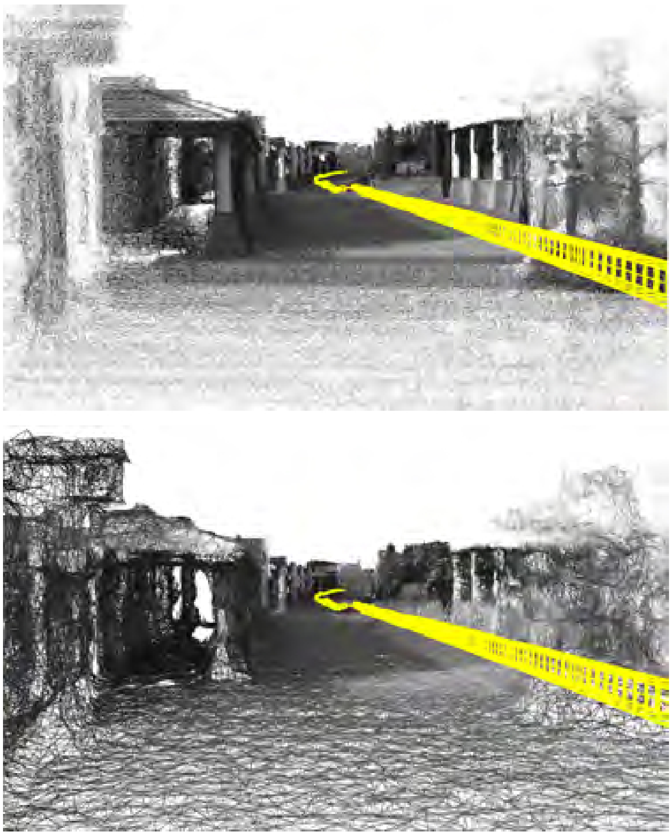

Figure 17: Reconstructed cloud of points (*left*) and surface (*right*) based on a single run of the ego-vehicle. Courtesy of Yi Zeng

6 Conclusions

The vehicle industry world-wide has assigned major research and development resources for offering competitive solutions for VB-DAS. Research at academic institutions needs to address future or fundamental tasks, challenges which are not of immediate interest for the vehicle industry, for being able to continue to contribute to this area. The chapter briefly reviewed work in the field of computer vision in vehicles.

Computer vision can help to solve true problems in society or industry, thus contributing to the prevention of social harms or atrocities. Academics identify *ethics in research* often with subjects such as plagiarism, competence, or objectivity, and a main principle is also social responsibility. Computer vision in road vehicles can play, for example, a major role in reducing casualties in traffic accidents which are counted by hundreds of thousands of people worldwide each year; it is a very satisfying task for a researcher to contribute to improved road safety. VB-DAS also contribute to a better driving comfort in modern cars. Autonomous driving is a realistic goal for some particular traffic situations in developed countries, but not yet expected as a general solution worldwide in the foreseeable future.

Acknowledgment. The author acknowledges comments made by Hui Chen, Clemens Dannheim, Uwe Franke, Antonio M. Lopez, Markus Mäder, Wolfgang Münst, Ralf Reulke, Junli Tao, Garry Tee, and Mohan M. Trivedi on drafts of this article.

References

[1] R. Klette. *Concise Computer Vision: An Introduction into Theory and Algorithms*. Springer, London, 2014.

[2] K. Bengler, K. Dietmayer, B. Farber, M. Maurer, C. Stiller, and H. Winner. Three decades of driver assistance systems: Review and future perspectives. IEEE Intelligent Transportation Systems Magazine, **6**:6–22, 2014.

[3] R. K. Jurgen (editor). *Adaptive Cruise Control*. SAE International, Warrendale, 2006.

[4] S. Motoyama, T. Ohta, T. Watanabe, and Y. Ito. Development of lane departure warning system. In Proc. *World Congress Intelligent Systems*, 8 pages, 2000.

[5] M. Á. Sotelo and J. Barriga. Blind spot detection using vision for automotive applications. J. Zhejiang University, Science A, **9**:1369–1372, 2008.

[6] J. Barnes. Daimler's double take sees machine vision move in-vehicle. *ITS International*, Nov./Dec. 2013.

[7] G. P. Stein, Y. Gdalyahu, and A. Shashua. Stereo-Assist: Top-down stereo for driver assistance systems. In Proc. *IEEE Intelligent Vehicles Symp.*, pages 723–730, 2010.

[8] WHO 2013 Global status report on road safety (online).

[9] P. Corke, J. Lobo, and J. Dias. An introduction to inertial and visual sensing. *Int. J. Robotics Research*, **26**:519–535, 2007.

[10] M. M. Trivedi, S. Y. Cheng, E. Childers, and S. Krotosky. Occupant posture analysis with stereo and thermal infrared video: Algorithms and experimental evaluation. IEEE Trans. Vehicular Technology, **53**: 1698–1712, 2004.

[11] R. Klette, N. Krüger, T. Vaudrey, K. Pauwels, M. van Hulle, S. Morales, F. Kandil, R. Haeusler, N. Pugeault, C. Rabe, and M. Lappe: Performance of correspondence algorithms in vision-based driver assistance using an online image sequence database. *IEEE Trans. Vehicular Technology*, **60**:2012–2026, 2011.

[12] A. Briassouli and I. Kompatsiaris. Change detection for temporal texture in the Fourier domain. In *Proc. ACCV*, LNCS 6492, pp. 149–160, 2010.

[13] B.-S. Shin, D. Caudillo, and R. Klette. Evaluation of two stereo matchers on long real-world video sequences. *Pattern Recognition*, DOI: 10.1016/j.patcog.2014.04.011, 2014.

[14] The *.enpeda..* Image Sequence Analysis Test Site, `ccv.wordpress.fos.auckland.ac.nz/eisats/`.

[15] The KITTI Vision Benchmark Suite. `www.cvlibs.net/datasets/kitti/`.

[16] M. Pavlic, G. Rigoll, and S. Ilic. Classification of images in fog and fog-free scenes for use in vehicles. In Proc. *IEEE Intelligent Vehicles Symposium*, pp. 481–486, 2013.

[17] 6D Vision. `www.6d-vision.com`.

[18] S. Morales and R. Klette. A third eye for performance evaluation in stereo sequence analysis. In Proc. *CAIP*, LNCS 5702, pp. 1078–1086, 2009.

[19] M. Rezaei and R. Klette. Look at the driver, look at the road: No distraction! No accident! In Proc. *CVPR*, pages 129–134, 2014.

[20] X. Jie, H. Chen, W. Ding, C. Zhao, and J. Morris. Robust optical flow for driver assistance. In Proc. *IVCNZ*, pp. 1-7, 2010.

[21] T. P. Michalke, F. Stein, and U. Franke. Towards a closer fusion of active and passive safety: Optical flow-based detection of vehicle side collisions. In Proc. *IEEE Intelligent Vehicle Symp.*, pp. 181–188, 2011.

[22] D. Ponsa, A. M. Lopez, J. Serrat, F. Lumbreras, and T. Graf. Multiple vehicle 3D tracking using an unscented Kalman filter. *Intelligent Transportation Systems Conf.*, 2005.

[23] Z. Zhang and Y. Shan. Incremental motion estimation through local bundle adjustment. 2001. MSR-TR-01-54, Microsoft Research, 2001.

[24] B. Triggs, P. McLauchlan, R. Hartley, and A. Fitzgibbon. Bundle adjustment - A modern synthesis. In *Proc. Vision Algorithms Theory Practice*, pp. 298–375, 2000.

[25] L. Kneip, M. Chli, and R. Siegwart. Robust real-time visual odometry with a single camera and an IMU. In Proc. *British Machine Vision Conf.*, pp. 16.1-16.11, 2011.

[26] S. Sivaraman and M. M. Trivedi. Looking at vehicles on the road: A survey of vision-based vehicle detection, tracking and behavior analysis. *IEEE Trans. Intelligent Transportation Systems*, **14**:1773–1795, 2013.

[27] H. Badino, U. Franke, and D. Pfeiffer. The stixel world - A compact medium level representation of the 3D-world. In Proc. *DAGM - Pattern Recognition*, pp. 51–60, 2009.

[28] Y. Lin, F. Guo, and S. Li. Road obstacle detection in stereo vision based on uv-disparity. *J. Information Computational Science*, **11**:1137–1144, 2014.

[29] X. Wang, M. Yang, S. Zhou, and Y. Lin. Regionlets for generic object detection. In Proc. *IEEE Int. Conf. Computer Vision*, pp. 17–24, 2013.

[30] C. Long, X. Wang, G. Hua, M. Yang, and Y. Lin. Accurate object detection with location relaxation and regionlets relocalization. In Proc. *Asian Conf. Computer Vision*, 2014.

[31] D. Ponsa, A. M. Lopez, F. Lumbreras, J. Serrat, and T. Graf. 3D vehicle sensor based on monocular vision. In Proc. *Intelligent Transportation Systems Conf.*, 2005.

[32] A. Ali and S. Afghani. Shadow based on-road vehicle detection and verification using Haar wavelet packet transform. In *Proc. IEEE Int. Conf. Information Communication Technologies*, pp. 346–350, 2005.

[33] D. Geronimo and A. M. Lopez. *Vision-based Pedestrian Protection Systems for Intelligent Vehicles*. Springer Briefs in Computer Science, 2013.

[34] J. Marin, D. Vazquez, A. M. Lopez, J. Amores, and B. Leibe. Random forests of local experts for pedestrian detection. In Proc. *Int. Conf. Computer Vision*, 2013.

[35] CVC-07 DPM Virtual-World Pedestrian Dataset. `www.cvc.uab.es/adas/site/?q=node/7`.

[36] T. Gandhi and M. M. Trivedi. Pedestrian protection systems: Issues, survey, and challenges. *IEEE Trans. Intelligent Transportation Systems*, **8**:413–430, 2007.

[37] A. Barth. Vehicle tracking and motion estimation based on stereo vision sequences. PhD thesis, Bonn University, 2010.

[38] E. Ohn-Bar and M. Trivedi. Fast and robust object detection using visual subcategories. In Proc. *CVPR Workshop Mobile Vision*, 2014.

[39] P. Felzenszwalb, R. Girshick, D. McAllester, and D. Ramanan. Object detection with discriminatively trained part-based models. *IEEE Trans. PAMI*, **32**:1627–1645, 2010.

[40] J.J. Yebes, L.M. Bergasa, R. Arroyo, and A. Lazaro. Supervised learning and evaluation of KITTI's cars detector with DPM. In Proc. *IEEE Intelligent Vehicle Symp.*, pp. 768–773, 2014.

[41] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In *Proc. CVPR*, pp. 886–893, 2005.

[42] L. Breiman. Random forests. *Machine Learning* **45**:5–32, 2001.

[43] J. M. Alvarez, A. M. Lopez, T. Gevers, and F. Lumbreras. Combining priors, appearance and context for road detection. *IEEE Trans. Intelligent Transportation Systems*, **15**:1168–1178, 2014.

[44] A. Wedel, U. Franke, H. Badino, and D. Cremers. B-spline modeling of road surfaces for freespace estimation. In Proc. *IEEE Intelligent Vehicles Symposium*, pp. 828–833, 2008.

[45] A. Wedel, H. Badino, C. Rabe, H. Loose, U. Franke, and D. Cremers. B-spline modeling of road surfaces with an application to free-space estimation. *IEEE Trans. Intelligent Transportation Systems*, **10**:572–583, 2009.

[46] B.-S. Shin, Z. Xu, and R. Klette. Visual lane analysis and higher-order tasks: A concise review. *Machine Vision Applications*, **25**:1519–1547, 2014.

[47] A. Bar Hillel, R. Lerner, D. Levi, and G. Raz. Recent progress in road and lane detection: a survey. *Machine Vision Applications*, **23**, published online, 2014.

[48] Z. Kim. Robust lane detection and tracking in challenging scenarios. *IEEE Trans. Intelligent Transportation Systems*, **9**:16–26, 2008.

[49] J. C. McCall and M. M. Trivedi. Video-based lane estimation and tracking for driver assistance: Survey, system, and evaluation. *IEEE Trans. Intelligent Transportation Systems*, **7**:20–37, 2006.

[50] A. M. Lopez, J. Serrat, C. Canero, F. Lumbreras, and T. Graf. Robust lane markings detection and road geometry computation. *Int. J. Automotive Technology*, **11**:395–407, 2010.

[51] Synthetic lane data. `www.cvc.uab.es/adas/projects/lanemarkings/IJAT/videos.html`.

[52] A. Borkar, M. Hayes, and M. T. Smith. An efficient method to generate ground truth for evaluating lane detection systems. In Proc. *IEEE Int. Conf. Acoustics Speech Signal Processing*, pp. 1090–1093, 2010.

[53] A. Møgelmose, M. M. Trivedi, and T. B. Moeslund. Vision based traffic sign detection and analysis for intelligent driver assistance systems: Perspectives and survey. *IEEE Trans. Intelligent Transportation Systems*, **13**:1484–1497, 2012.

[54] S. Escalera, X. Baró, O. Pujol, J. Vitrià, and P. Radeva. *Traffic-Sign Recognition Systems*. Springer, 2011.

[55] N. Barnes and A. Zelinsky. Real-time radial symmetry for speed sign detection. In Proc. *IEEE Intelligent Vehicles Symposium*, pp. 566–571, 2004.

[56] S. Müller-Schneiders, C. Nunn, and M. Meuter. Performance evaluation of a real time traffic sign recognition system. In Proc. *IEEE Intelligent Vehicles Symposium*, pp. 79–84, 2008.

[57] J. Crisman and C. Thorpe. Unscarf: A color vision system for the detection of unstructured roads. In Proc. *IEEE Int. Conf. Robotics Automation*, volume 3, pp. 2496–2501, 1991.

[58] M. Schreier and V. Willert. Robust free space detection in occupancy grid maps by methods of image analysis and dynamic B-spline contour tracking. In Proc. *IEEE Conf. Intelligent Transportation Systems*, pp. 514 – 521, 2012.

[59] S.-J. Jung, H.-S. Shin, and W.-Y. Chung. Driver fatigue and drowsiness monitoring system with embedded electrocardiogram sensor on steering wheel. *IET Intelligent Transport Systems*, **8**:43–50, 2014.

[60] P. Viola and M. Jones. Robust real-time face detection. *Int. J. Computer Vision*, **57**:137–154, 2001.

[61] E. Murphy-Chutorian and M. M. Trivedi. Head pose estimation in computer vision: A survey. *IEEE Trans. Pattern Analysis Machine Intelligence*, **31**:607–626, 2009.

[62] A. Tawari, K. H. Chen, and M. M. Trivedi. Where is the driver looking: Analysis of head, eye and iris for robust gaze zone estimation. In Proc. *IEEE Intelligent Transportation Systems Conf.*, pp. 988–994, 2014.

[63] M. Rezaei and R. Klette. Novel adaptive eye detection and tracking for challenging lighting conditions. In Proc. *ACCV Workshops*, LNCS 7729, pp. 427–440, 2012.

[64] C. Tran, A. Doshi, and M. M. Trivedi. Modeling and prediction of driver behavior by foot gesture analysis. *Computer Vision Image Understanding*, **116**: 435–445, 2012.

[65] E. Ohn-Bar and M. M. Trivedi. Beyond just keeping hands on the wheel: Towards visual interpretation of driver and motion patterns. In Proc. *IEEE Intelligent Transportation Systems Conf.*, pp. 1245–1250, 2014.

[66] M. L. Eichner and T. P. Breckon. Integrated speed limit detection and recognition from real-time video. In Proc. *IEEE Intelligent Vehicles Symposium*, pp. 626–631, 2008.

[67] A. Kheyrollahi and T.P. Breckon. Automatic real-time road marking recognition using a feature driven approach. *Machine Vision Applications*, **23**:123–133, 2012.

[68] G. Grubb, E. Jakobsson, A. Beutner, M. Ahrholdt, and S. Bergqvist. Automatic queue assistance to aid under-loaded drivers. In CD *ITS World Congress Exhibition Intelligent Transport Systems Services*, 8 pages, 2009.

[69] M. Yu and G. Ma. 360° surround view system with parking guidance. *SAE Int. J. Commer. Veh.*, **7**:19–24, 2014.

[70] C. Wang, H. Zhang, M. Yang, X. Wang, L. Ye, and C. Guo. Automatic parking based on a bird's eye view vision system. *Advances in Mechanical Engineering*, Article ID 847406, 13 pages, 2014.

[71] R. Preiß, C. Gruner, T. Schilling, and H. Winter. Mobile vision - Developing and testing of visual sensors for driver assistance systems. In *Advanced Microsystems for Automotive Applications* (J. Valldorf and W. Gessner, editors), VDI-Buch, pp. 95–107, 2004.

[72] Technology award for BLIS in 2006. `www.roadsafetyawards.com/national/view.aspx?winnerid=107`.

[73] D. Bétaille, R. Toledo-Moreo, and J. Laneurit. Making an enhanced map for lane location based services. In Proc. *IEEE Conf. Intelligent Transportation Systems*, pp. 711–716, 2008.

[74] F. Peyret, J. Laneurit, and D. Bétaille. A novel system using enhanced digital maps and WAAS for a lane-level positioning. In Proc. *World Congress Intelligent Transport Systems*, 12 pages, 2008.

[75] Y. Sekimoto, Y. Matsubayashi, H. Yamada, R. Imai, T. Usui, and H. Kanasugi. Light weight lane positioning of vehicles using a smartphone GPS by monitoring the distance from the center line. In Proc. *IEEE Conf. Intelligent Transportation Systems*, pp. 1561–1565, 2012.

[76] A. Selloum, D. Bétaille, E. Le Carpentier, and F. Peyret. Robustification of a map aided location process using road direction. In Proc. *IEEE Conf. Intelligent Transportation Systems*, pp. 1504–1510, 2010.

[77] H. Deusch, J. Wiest, S. Reuter, M. Szczot, M. Konrad, and K. Dietmayer. A random finite set approach to multiple lane detection. In Proc. *IEEE Conf. Intelligent Transportation Systems* pp. 270–275, 2012.

[78] R. A. Gupta, W. Snyder, and W. S. Pitts. Concurrent visual multiple lane detection for autonomous vehicles. In Proc. *IEEE Conf. Robotics Automation*, pp. 2416–2422, 2010.

[79] K. Zhao, M. Meuter, C. Nunn, D. Muller, S. Muller-Schneiders, and J. Pauli. A novel multi-lane detection and tracking system. In Proc. *IEEE Intelligent Vehicles Symposium*, pp. 1084–1089, 2012.

[80] J. Tao, B.-S. Shin, and R. Klette. Wrong roadway detection for multi-lane roads. In Proc. *CAIP*, LNCS 8048, pp. 50–58, 2013.

[81] Daimler A.G. First assistance system in the fight against vehicles driving on the wrong side of the road. `www.daimler.com/dccom/`, January 21, 2013.

[82] G. Monteiro, M. Ribeiro, J. Marcos, and J. Batista. Wrongway drivers detection based on optical flow. In Proc. *IEEE Int. Conf. Image Processing*, volume 5, pp. 141–144, 2007.

[83] J. C. Rubio, J. Serrat, A. M. Lopez, and D. Ponsa. Multiple target tracking for intelligent headlights control. *IEEE Trans. Intelligent Transportation Systems*, **13**:594–609, 2012.

[84] M. M. Trivedi, T. Gandhi, and J. McCall. Looking-in and looking-out of a vehicle: Computer-vision-based enhanced vehicle safety. *IEEE Trans. Intelligent Transportation Systems*, 2007.

[85] J. C. McCall and M. M. Trivedi. Driver behavior and situation aware brake assistance for intelligent vehicles. *Proc. IEEE*, 2007.

[86] A. Tawari, S. Sivaraman, M. M. Trivedi, T. Shanon, and M. Tippelhofer. Looking-in and looking-out vision for urban intelligent assistance: Estimation of driver attention and dynamic surround for safe merging and braking. In Proc. *IEEE Intelligent Vehicles Symp.*, 2014.

[87] A. Doshi, B. Morris, and M. Trivedi. On-road prediction of driver's intent with multimodal sensory cues. *IEEE Pervasive Computing*, 2011.

[88] E. Ohn-Bar, A. Tawari, S. Martin, and M. M. Trivedi. On surveillance for safety critical events: In-vehicle video networks for predictive driver assistance systems. *Computer Vision Image Understanding*, 2015.

[89] G. Ros, A. Sappa, D. Ponsa, and A. M. Lopez. Visual SLAM for driverless cars: a brief survey. In Proc. *IV Workshop Navigation Perception Accurate Positioning Mapping Intelligent Vehicles*, 2012.

[90] G. Ros, S. Ramos, M. Granados, A. H. Bakhtiary, D. Vazquez, and A. M. Lopez. Vision-based offline-online paradigm for autonomous driving. In Proc. *Winter Conf. Applications Computer Vision*, 2015.

[91] U. Franke, D. Pfeiffer, C. Rabe, C. Knoeppel, M. Enzweiler, F. Stein, and R. G. Herrtwich. Making Bertha see. In Proc. *IEEE Int. Conf. Computer Vision Workshops*, pp. 214–221, 2013.

[92] M. Debord. 3 reasons why the auto industry should be terrified of the Google car. *Business Insider Australia*, 27 Dec 2014.

[93] J. Xiao, T. Fang, P. Zhao, M. Lhuilier, and L. Quan. Image-based streetside city modeling. In Proc. *SIGGRAPH*, article 114, 2009.

[94] A. Geiger, J. Ziegler, and C. Stiller. StereoScan: Dense 3d reconstruction in real-time. In Proc. *IEEE Intelligent Vehicles*, pp. 963–968, 2011.

[95] F. Huang and R. Klette. City-scale modeling towards street navigation applications. *J. Information Convergence Communication Engineering*, **10**, 411–419, 2012.

[96] Y. Zeng and R. Klette. Multi-run 3D streetside reconstruction from a vehicle. In *Proc. CAIP*, LNCS 8047, pp. 580–588, 2013.