COMPUTER-AIDED DESIGN & APPLICATIONS, 2018
VOL. 15, NO. 4, 501–508
https://doi.org/10.1080/16864360.2017.1419638

Sensor and object recognition technologies for self-driving cars

Mario Hirz and Bernhard Walzel

Graz University of Technology, Austria

ABSTRACT
Autonomous driving functions for motorized road vehicles represent an important area of research today. In this context, sensor and object recognition technologies for self-driving cars have to fulfill enhanced requirements in terms of accuracy, unambiguousness, robustness, space demand and of course costs. The paper introduces different levels of automated driving functions according to the SAE standard and derives requirements on sensor technologies. Subsequently, state of the art technologies for object detection and identification as well as systems under development are introduced, discussed and evaluated in view of their suitability for automotive applications.

KEYWORDS
Autonomous driving; sensor technology; object recognition; image-based geometry modeling; object differentiation

1. Introduction

Assisted and automated driving technologies are one of the most intensive researched and investigated fields in automotive industry today. Nearly all car manufacturer and several well-known player from the electronics and communication industry push forward knowhow for self-driving cars. Scientific and non-scientific community intensively discusses the introduction of automated driving functions in cars on worldwide markets as well as the potentials, benefits and risks of this technology for humans and environment.

In this way, it is expected that vehicles with driving assistance functions and even autonomous vehicles for passenger and goods transportation will have an increasingly share in daily traffic [11], [20]. This requires an integration of diverse types of technologies into existing traffic systems. According to SAE [17] there are different levels of autonomous driving, Fig. 1. Level “0” represents the old-fashioned car that is thoroughly controlled by a human driver. Most of the cars today are in this level. With raising SAE level, the share of automated driving functions is increasing. Level “1” and “2” include driving assistance supporting the chauffeur in certain traffic situations, e.g. automated cruise control, lane assistant, highway assistant. In these technology levels, the car driver has to keep an eye on the driving situations continuously to be able to perform interventions. From level “3” on, the car takes over control in certain traffic conditions, e.g. highway chauffeur or traffic jam assistant. Here, the driver is able to do other things in the car, e.g. reading, writing or relaxing. But an important issue of highly automated driving functions in level “3” is that the driver has to take over control, if the automated functions are disturbed or not able to handle complex traffic situations. In level “4”, fully automated functions provide all required operations of driving control, so that no human interventions are required. The driver has the possibility to take over control and to steer the car by hand in case of system fail functions or personal desire, but in general the car acts as a self-driving device. Finally, level “5” represents the full autonomous vehicle that does not need human interaction anymore. In this level, the cars are robots that transport passengers and/or goods independently.

Besides technology-related research, self-driving cars are topic of far-reaching evaluations and investigations under consideration of e.g. legislative, social and ethic aspects. Pros and cons are discussed intensively, because of the big impact this technology will have on human mobility behavior, infrastructure and environment. This includes great potential to improve quality of life, reduce traffic accidents and decrease impacts on environmental pollution, but of course there are many open points in terms of reliability, costs, safety and ethical issues.

An important basis for the realization of autonomous driving function represents a reliable and robust determination of position and trajectory of the ego-vehicle on road for tracking the driving paths. In addition to its own position, the behavior of all other traffic participants has to be observed and predicted. This includes

CONTACT Mario Hirz mario.hirz@tugraz.at; Bernhard Walzel bernhard.walzel@tugraz.at

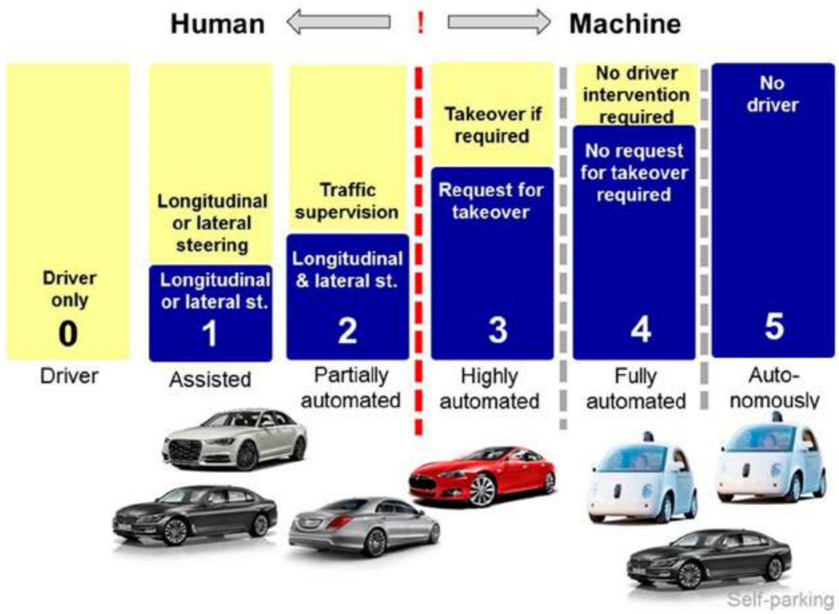

Figure 1. Levels of autonomous driving according to SAE J3016.

a safe detection of position and trajectories of other vehicles as well as the cognition of intentions and gestures of pedestrians. In combination with environmental models, e.g. digital maps, the recognition of all traffic participants is required for a continuously modeling of the traffic situation in different scenarios, e.g. on highway, country roads or in urban areas. In addition, specific scenarios have to be handled, e.g. in parking garages. Besides recognition and modeling of environment and traffic, car to car and car to infrastructure communication (car2x-communication) plays an important role. All the information is finally proceeded by real-time data fusion and evaluation [5], [3]. Periphery sensing functions include both, sensor data processing and fusion as well modeling and interpretation of traffic situations. Traffic situations include all influencing factors in a comprehensive way: ego-car, driving situation, traffic around the car, environmental boundaries and surrounding events, e.g. a group of pedestrians that is crossing the street.

2. Sensor technologies for autonomous driving

For detection of position, geometry, type and motion of objects, different types of sensors can be used, which all have their specific characteristics. Fig. 2 shows a general overview of technologies that can be used for obstacle recognition and data transfer including a comparison of operating range and accuracy.

Fig. 3 points to the technologies that are applicable for in-car sensor systems for environment and object recognition. Far-spread since years are ultrasonic sensors with an operating range of about 2 m, which are exemplary used for car parking assistant functions. In terms of image recognitions, camera systems come to use. This technology represents a cost effective way for object identification, but the sensing quality strongly depends on environmental aspects, e.g. weather and light reflection. Night view technology is based on infrared camera systems. Radar sensors are robust in terms of different weather conditions and they have a long operating range of up to 250 m, but the delivered information about obstacle geometry is limited. Long-range radio detection and ranging is used in automated cruise control functions to identify other cars on road. Front and rear end mid-range radar systems with a range from 30 to about 150 m are exemplary used in emergency brake assistants and high way assistants (supporting automated overtaking processes). Finally, laser scanner (LIDAR – light detection and ranging) represent a high sophisticated but expensive technology. By use of rotating laser light impulses, LIDAR are able to create detailed 3D-maps of the surrounding environment, considering an operating range of more than 100 m.

The mentioned sensor technologies are used according to their ranges of application and combined with global positioning systems (GPS). External information is delivered by car2x-communication, e.g. 3D maps, information regarding traffic situations, on-road construction areas or dangerous road and weather situations. The combination process of all these data delivered by sensors, positioning and communication systems

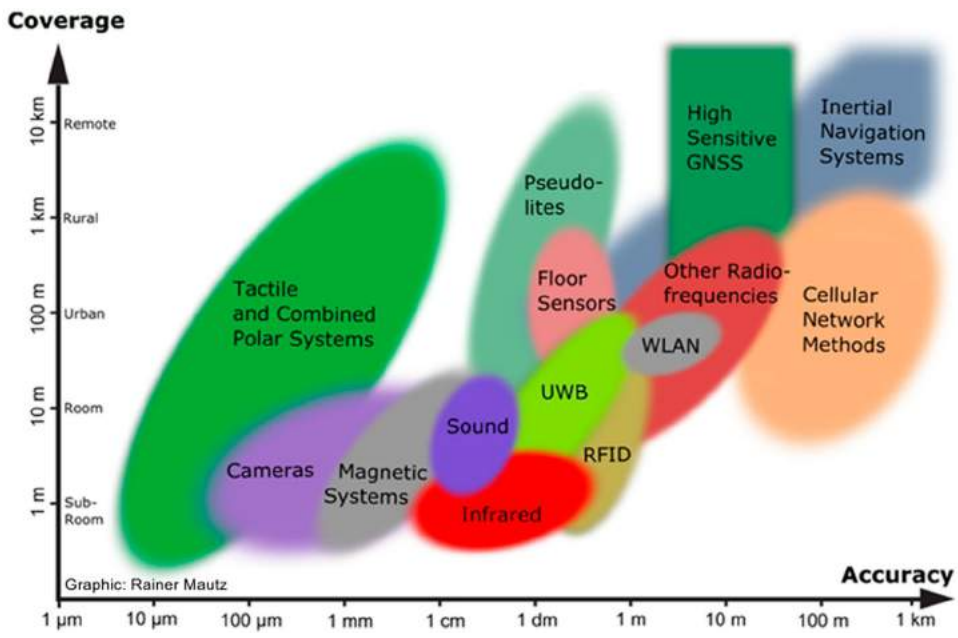

Figure 2. Operating range and accuracy of different sensor technologies [12].

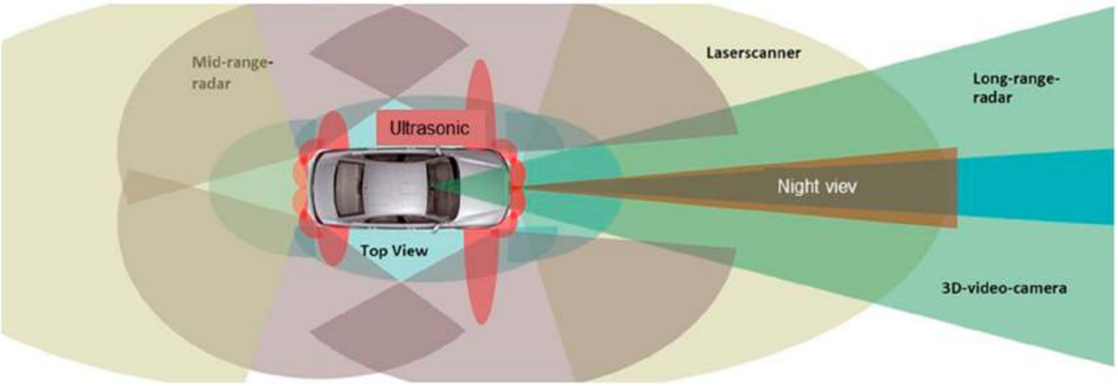

Figure 3. In-car sensor technologies for environment and object recognition, according to [18].

includes data association and the actual traffic situation estimation [9]. For the estimation of targets, the traffic situation is modeled and interpreted by use of geometrical models. These models serve as a basis for cognition algorithms that recognize and evaluate the driving situation in real time. Subsequently, planning, simulation and automated execution of driving maneuvers can be performed, while the above mentioned sensing and detection processes are running simultaneously. Due the considerable complexity of situation modeling and interpretation processes, artificial intelligence (AI) based approaches come to use. A key role plays object recognition and interpretation that runs in real-time. In this context, the mentioned sensor data are processed to generate geometry models of environment and traffic-related objects.

3. Different technologies for object recognition

For automated driving functions, three technologies are in use and in further development: image recognition with camera systems, radar detection and light detection & ranging. Due to its limited operating range, ultrasonic sensors are used for near field observation only. Whereby image recognition and radar systems are applied in mass production cars, LIDAR comes to use in research vehicles today. It provides great possibilities, but represents the most expensive technology.

3.1. Image recognition

Camera based systems are able to deliver detailed information of traffic situations, including objects on or

Figure 4. Application of camera systems for object recognition: left daylight camera, right infrared camera, [13], [8].

alongside roads, traffic signs, road markings and road conditions as well as surrounding objects, e.g. buildings. In this way, image recognition is important for traffic situation modeling, because the system is able to see the situation in the same way as a human driver. Today's systems in market-available cars work with one or more cameras placed at the front side, back side and wing mirrors of the vehicle. They are able to detect specific classes of objects, including people, traffic signs and cars within complex images by use of generative and discriminative statistical classification models. To enhance the possibilities of camera based systems for cases with reduced light intensity, some manufacturer established infrared night vision systems. Fig. 4, left, shows an exemplary application of the camera system that is used in the Tesla Model S to supply the autopilot function with required information. On the right side of Fig. 4, an application of an infrared camera system for human objects recognition is shown.

Basic functionality of computer vision, so-called machine vision, is used for detection of particular (predefined) features or object characteristics, e.g. edges, corners, as well as for motion detection and distance estimation. In this way it is used for the determination of lane markings, edges of roads, the general position of other vehicles or obstacles on and nearby the road [1], [14]. Enhanced functionalities of computer vision provide a more accurate recognition of objects, but require another, more complex approach that includes machine learning techniques to teach an AI to recognize and classify certain objects. Target of ongoing research is to train systems in a way that they are able to not only detect objects, but also to differentiate with a high accuracy. Exemplary, today's camera systems in cars are not able to distinguish between a little child and a dog. But in some dangerous traffic situations, it might be essential that the system has the information, if a little child or a dog is unexpectedly crossing the road.

Fig. 5 shows a comparison of different image-based object recognition technologies. A scaled car model serves as test object; in addition some markings have been applied on the parking surface of the car. Three different types of cameras come to use, whereby all cameras have been mounted directly above the test object with a distance of 1 m. Target is to evaluate sensor signal quality and geometrical modeling approaches under different boundary conditions, e.g. lighting intensity and direction, light reflection at diverse surfaces. In case of a 2D camera, the system delivers a 2D image without depth information. The sensor signal is processed in edge detection algorithms that are able to extract the outline shape of the car [19]. The detection of inside edges strongly depends on the lighting situation and on the type and quality of surfaces. Exemplary, the upper front screen edge is not recognized due to the reflection of light in this area. In this way, 2D-camera based systems are able to deliver information about object size and position, but only marginal data about class and type of the objects.

3D cameras are able to deliver information about the depth of field of an image. This technology is well known from gaming industry, e.g. [21], and is approaching industrial applications today. 3D information can be generated in different ways, e.g. by structured light [16], or by time-of-flight (ToF) technology [7]. Both methods are relatively simple and cost-efficient, which makes them interesting for large-scale applications in the automotive industry. The result of the structure light based object geometry rebuilding shows that a considerable share of the outer surfaces is not detected (white areas). This leads to missing geometry information and thus to a bad model quality. This technology is very sensitive in terms of light reflections. In addition, the geometrical recognition tolerance of inclined surfaces is coarse. ToF technology provides better geometrical recognition results due to the fact, that not only a 2D-image is captured, but also the distance from the camera to each pixel by use of a light transient time measurement. This technology is less sensitive in case of varying lighting conditions and reflections. The gathered sensor data are processed and converted into 3D geometry models that

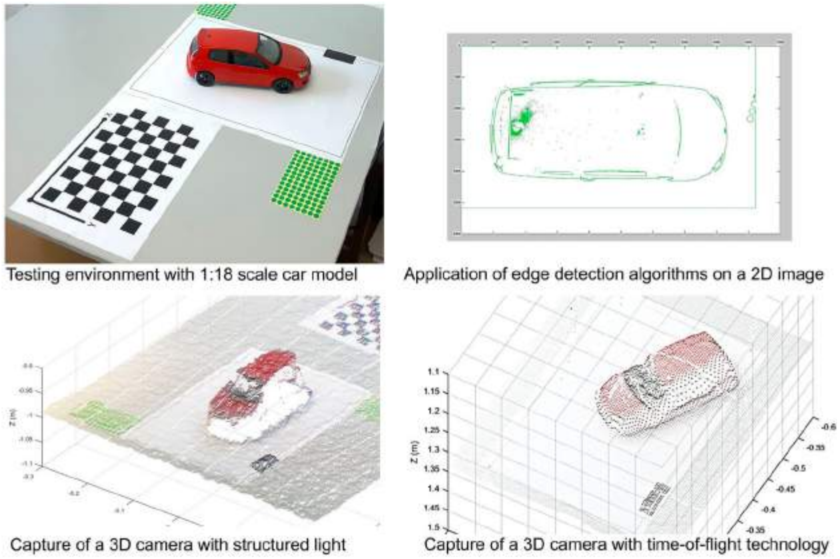

Figure 5. Different image-based object recognition technologies.

deliver information for object recognition algorithms. ToF camera systems have a working range of up to 40 m with an accuracy of about 0.01 m. The image refresh rates reach 160 captures per second, which allows real-time applications. Nowadays, ToF cameras come to use in driver assistance technologies, emergency brake systems and pedestrian protection. Research goes into the direction of a combination of 3D ToF technology with 2D image recognition to increase the accuracy of object identification and differentiation.

3.2. Radar

Radar sensors measure the distance and relative velocity to objects precisely and in a high accuracy. An important advantage of radar-based systems is their robust and reliably operation under different environmental conditions, e.g. weather, dust and pollution. Drawbacks include higher costs in comparison with camera-based systems and a relatively coarse representation of objects that is based on a limited number of radar lobes. In addition, identikit pictures occur in case of reflections and perturbations [6], [2].

Radar based systems are applied in automated cruise control, emergency brake systems and collision detection as well as in lane change assistance and blind spot detection systems. Due to the fact, that radar sensors are not able to deliver detailed image- and geometry information of objects, the application for object recognition purposes is limited.

In this way, radar sensor systems are often combined with camera based image recognition, so that each technology plays its advantages. Exemplary, radar sensors are used for obstacle detection and relative velocity measurement, whereby the combined 2D-camerea system delivers information of type and class of objects. This is exemplary used for pedestrian recognition, traffic sign identification or traffic lane cognition. Fig. 6 shows the results of sensor data processing of a serial automotive radar system. The right image shows a driving situation on highway and the left diagram the modeled objects. In the diagram, the ego-car is placed at position (0; 0) and all detected objects in front of the ego-car are represented as rectangles, e.g. the car driving ahead is shown as purple rectangle. It is visible, that the system detects lots of objects, but most of them are aside or above the road. This requires a complex sensor data processing to provide sufficient object classification and identification.

3.3. LIDAR

Light detection and ranging is seen as the most powerful, but most expensive technology for object recognition. A fast rotating laser sensor provides several millions of data points per second, enabling the creation of detailed 3D maps of both surrounding objects and environment. Especially the mapping of environment including area topology, buildings and objects aside the road is important for autonomous driving functions. In this way, highly accurate maps of the infrastructure are generated and provided to the vehicle [13]. In Fig. 7, an actual point cloud delivered by a LIDAR represents a traffic situation at an intersection including other vehicles and road features around.

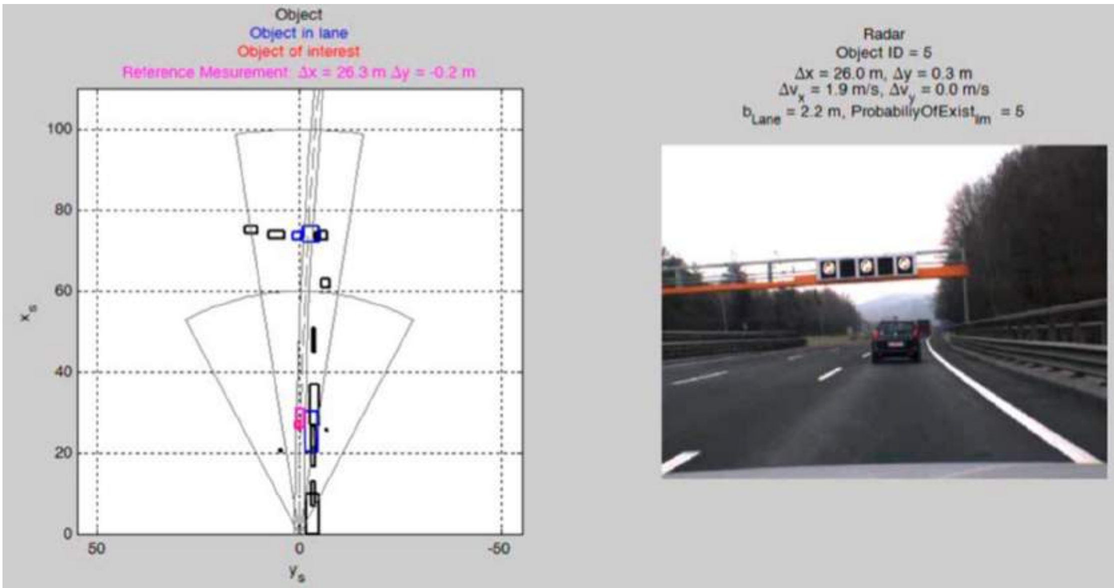

Figure 6. Object modeling based on automotive radar sensor signals [4].

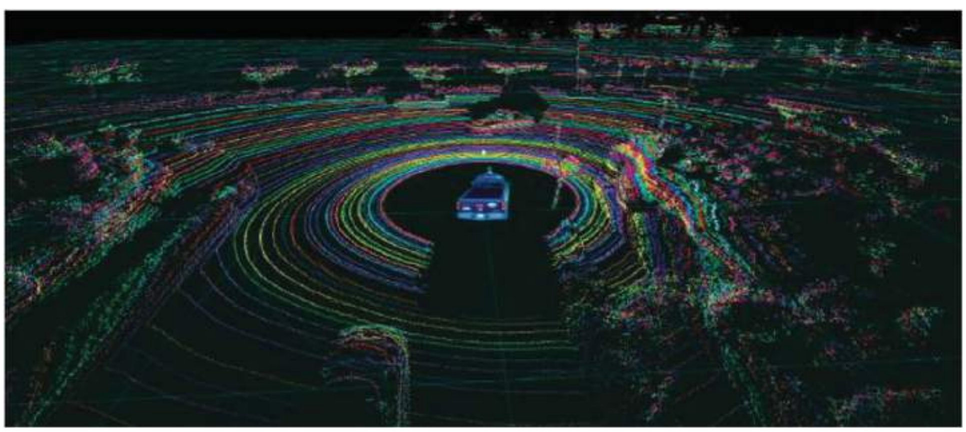

Figure 7. Point cloud of a traffic situation delivered by LIDAR [10].

A key advantage of the technology in comparison with camera based system is that the function is not influenced by ambient lighting conditions, because LIDAR emits laser light. On the other hand, cameras provide a higher resolution and are able to detect colors. With several ten thousand dollars, LIDAR systems are very expensive today, but there are new technologies approaches that are able to decrease the costs significantly, e.g. stationary LIDAR without moving parts inside. For self-driving cars, a combination of LIDAR with camera based systems promise the largest potential in terms of object and environment recognition and differentiation. LIDAR is suitable for accurate determination of object positions and trajectories as well as for environment mapping; computer vision technology is advantageous for object recognition and identification as well as for detailed image processing, e.g. face detection.

4. Comparison and evaluation

In Fig. 8 a summary and evaluation of the mentioned sensor technologies for automated driving functions is given. The evaluation is performed based on selected test vehicles, equipped with the selected sensor systems. As one important criteria, the sensitiveness in view of different ambient light conditions is investigated and the robustness in view of weather conditions is considered. Besides sensing accuracy and object recognition capability (e.g. for the detection and classification of other vehicles on

Sensor	Bright light performance	Low light performance	Outdoor	Weather robustness	Vehicle classification	Vehicle adaptation	Material costs
Ultrasonic	Good	Good	Yes	Good	No	No	Low
Magnetic	Good	Good	Yes	Good	No	Yes	High
2D-camera	Good	Weak	Yes	Weak	Yes	No	Low
Laser and lidar	Good	Good	Yes	Good	Yes	No	Very High
3D-imaging technologies							
3D-camera	Good	Weak	Yes	Weak	Yes	No	Low
Structured light	Weak	Good	No	Weak	Yes	No	Medium
ToF-cameras	Good	Good	Yes	Good	Yes	No	High

Figure 8. Comparison of different sensor technologies for automated driving functions.

road), the effort for vehicle adaptation and costs is also considered.

5. Conclusion

Object recognition and differentiation delivers important information for automated driving functions. It requires both enhanced sensor technology and object- respectively geometry modeling. Camera based technologies provide detailed information about object surface and appearance characteristics, but have limited performance under bad weather and low light conditions and in terms of position, velocity and driving trajectory determination. Radar systems deliver reliable information about object position and velocity even under bad environmental conditions, but due to the applied technology they are not able to provide data about object type, geometry and surface information. Finally, laser sensors are characterized by an enormous captured data volume that includes detailed information about the distance, velocity and trajectory of moving objects as well as detailed geometry data of the environment and surrounding items. The only thing, laser sensors cannot recognize information on surfaces, e.g. written letters on a traffic sign. The trend in object recognition technologies for self-driving cars goes into the direction of sensor fusion that combines camera based technologies with radar or laser sensors, so that each technology is able to play its advantages in a highly integrated system.

Acknowledgement

This work was funded by the Austrian Research Funding Association (FFG) under the scope of the Mobility of the Future Program.

ORCID

Mario Hirz http://orcid.org/0000-0002-4502-4255
Bernhard Walzel http://orcid.org/0000-0003-4818-3726

References

[1] Alkhorshid, Y.; et al.: Camera-Based Lane Marking Detection for ADAS and Autonomous Driving, International Conference Image Analysis and Recognition – ICIAR 2015, Springer, doi:10.1007/978-3-319-20801-5_57
[2] Amsel, Ch.; Stapel, D.: Compact radar sensors enabling autonomous driving, 17. Internationales Stuttgarter Symposium, Germany, 2017, doi:10.1007/978-3-658-16988-6_67
[3] Ayşegül, U.; Yakup D.; Cüneyt G.: Object recognition and detection with deep learning for autonomous driving applications, SAGE Journal, 2017, doi:10.1177/0037549717709932
[4] Bernsteiner, S.: Integration of Advanced Driver Assistance Systems on Full-Vehicle Level, PhD-Thesis at Graz University of Technology, Austria, 2016, ISBN: 978-3-85125-468-6
[5] Cheng, H.: Autonomous Intelligent Vehicles – Theory, Algorithms, and Implementation, 2001, Springer, ISBN 978-1-4471-2280-7
[6] Greco, M.S.: Automotive Radar, 2012 IEEE Radar Conference, Atlanta, USA
[7] Hansard, M.; Lee, S.; Choi, O.; Horaud, R.: Time of Flight Cameras: Principles, Methods, and Applications, Springer, 2012, ISBN 978-1-4471-4658-2
[8] Infrared Night Vision System, Guideinfrared, http://guideinfrared.com, [20170112]
[9] Lee, S.; Ko, H.; Hahn, H.: Multisensor Fusion and Integration for Intelligent Systems, Springer, 2009, ISBN 978-3-540-89858-0, doi:10.1007/978-3-540-89859-7
[10] LIDAR Sensor mapping, VelodyneLiDAR, CA, USA, 2017, http://velodynelidar.com, [20170112]
[11] Maurer, M.; et al.: Autonomous Driving – Technical, Legal and Social Aspects, open access book at link.springer.com, ISBN 978-3-662-48847-8, doi:10.1007/978-3-662-48847-8_1
[12] Mautz R.: Indoor positioning technologies, Institute of Geodesy and Photogrammetry, Department of Civil, Environmental and Geomatic Engineering, ETH Zürich, Switzerland, 2012
[13] Min, K.; Choi, J.: Vehicle Positioning Technology Using Infra-based Laser Scanner Sensors for Autonomous Driving Service, Computer Science and Convergence CSA

2011 & WCC 2011 Proceedings, Springer, ISBN: 978-94-007-2791-5
[14] Petrovskaya, A.; Thrun, S.: Model based vehicle detection and tracking for autonomous urban driving, Autonomous Robots vol. 26, 2009, Springer, doi:10.1007/s10514-009-9115-1
[15] Pirzada, U.: Tesla's Autopilot System – MobilEye, Wccftech, 2015, http://wccftech.com, [20170112]
[16] Rocchini, C.; Cignoni, P.; Montani, C.; Pingi, P.; Scopigno, R.: A low cost 3D scanner based on structured light, EUROGRAPHICS 2001, Oxford, USA, 2001.
[17] SAE Standard J3016: Taxonomy and Definitions for Terms Related to On-Road Motor Vehicle Automated Driving Systems, 2014, USA, www.sae.org/autodrive, [20161012]
[18] Sensors for autonomous cars, Audi Media Centre: https://www.audi-mediacenter.com, [20151010]
[19] Szeliski, R.: Computer Vision: Algorithms and Applications, Springer, 2010, ISSN 1868-0941
[20] Watzenig, D.; Horn, M.: Automated Driving – Safer and More Efficient Future Driving, Springer, 2017, ISBN 978-3-319-31895-0, doi:10.1007/978-3-319-31895-0_1
[21] Zeng, W.: Microsoft Kinect Sensor and Its Effect, IEEE Multimedia, Univ. of Missouri, 2012.