

Intelligent In-Vehicle Interaction Technologies

Prajval Kumar Murali, Mohsen Kaboli,* and Ravinder Dahiya*

With rapid advances in the field of autonomous vehicles (AVs), the ways in which human–vehicle interaction (HVI) will take place inside the vehicle have attracted major interest and, as a result, intelligent interiors are being explored to improve the user experience, acceptance, and trust. This is also fueled by parallel research in areas such as perception and control of robots, safe human–robot interaction, wearable systems, and the underpinning flexible/printed electronics technologies. Some of these are being routed to AVs. Growing number of network of sensors are being integrated into the vehicles for multimodal interaction to draw correct inferences of the communicative cues from the user and to vary the interaction dynamics depending on the cognitive state of the user and contextual driving scenario. In response to this growing trend, this timely article presents a comprehensive review of the technologies that are being used or developed to perceive user's intentions for natural and intuitive in-vehicle interaction. The challenges that are needed to be overcome to attain truly interactive AVs and their potential solutions are discussed along with various new avenues for future research.

(OEM) such as Tesla, BMW, Daimler, and big technology giants such as Google's Waymo and Apple Car pushing toward the full self-driving objective. According to the SAE J3016 taxonomy^[1] of autonomous driving, there are six levels of automated driving systems ranging from level 0 (completely manual) to level 5 (full self-driving [FSD]) automated systems that are expected to function in all geographic locations, all weather conditions, and under all conditions. The proposed benefits of intelligent vehicles include fewer road accidents, enhanced safety, reduced traffic congestion, effective commute time usage, and importantly enjoyable and comfortable ride. With increased autonomy, the drivers also assume the role of mere passengers who are engaged in nondriving activities and are unavailable to participate in traffic interactions. This would increase complexity in a mixed-autonomy traffic environ-

ment as interactions with pedestrians and cyclists are based on visual cues with the driver. Therefore, intelligent vehicles also need to autonomously interact with other traffic participants such as pedestrians and cyclists and other vehicles.

Human–vehicle interaction (HVI) is closely related to the field of human–robot interaction (HRI). It entails the problem of understanding and shaping the interaction dynamics between humans and vehicles. Particularly, the domain of interaction involves the study of sensation, perception, information exchange, inference, and decision-making.^[2] With increasing levels of automation, the driver will have more time and choice to perform various tasks other than driving and this opens new avenues for interaction. As a result, a growing number of sensors are being incorporated into the vehicles to understand the driver's and/or passenger's actions, emotions, and personal choices to offer the precise functionalities and services for an enjoyable ride.^[3] **Figure 1** shows the schematic of in-vehicle interaction consisting of various interactive interfaces and sensing modalities present in the vehicle as well as some modes of interaction such as gestures, speech, and eye gaze.

The advances in HVI have been presented in previous review articles, which have mainly focused on the detection of particular characteristics for driving assistance such as driver attention detection,^[4] emotion recognition,^[5] drowsiness detection,^[6] mental workload,^[7] and so on. A few other articles have reviewed the interior designs of AVs to better support and interact with drivers and passengers.^[8,9] Likewise, the user interfaces (UIs) and user experience (UX) of vehicle interiors have been studied in context with manual as well as automated driving.^[10] A number of previous works have also studied the interaction with the external


1. Introduction

Autonomous vehicles (AVs) will soon become a reality with almost every major automotive original equipment manufacturer

P. K. Murali, M. Kaboli
BMW Group
Department of Research
New Technologies
Innovations
Parkring 19, 85748 Garching bei München, Germany
E-mail: Mohsen.Kaboli@bmwgroup.com

P. K. Murali, R. Dahiya
Bendable Electronics and Sensing Technologies (BEST) Group
James Watt School of Engineering
University of Glasgow
G12 8QQ Glasgow, UK
E-mail: Ravinder.Dahiya@glasgow.ac.uk

M. Kaboli
Donders Institute for Brain
Cognition and Behaviour
Radboud University
Heyendaalseweg 135, 6525 AJ Nijmegen, The Netherlands

 The ORCID identification number(s) for the author(s) of this article can be found under <https://doi.org/10.1002/aisy.202100122>.

© 2021 The Authors. Advanced Intelligent Systems published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

DOI: 10.1002/aisy.202100122

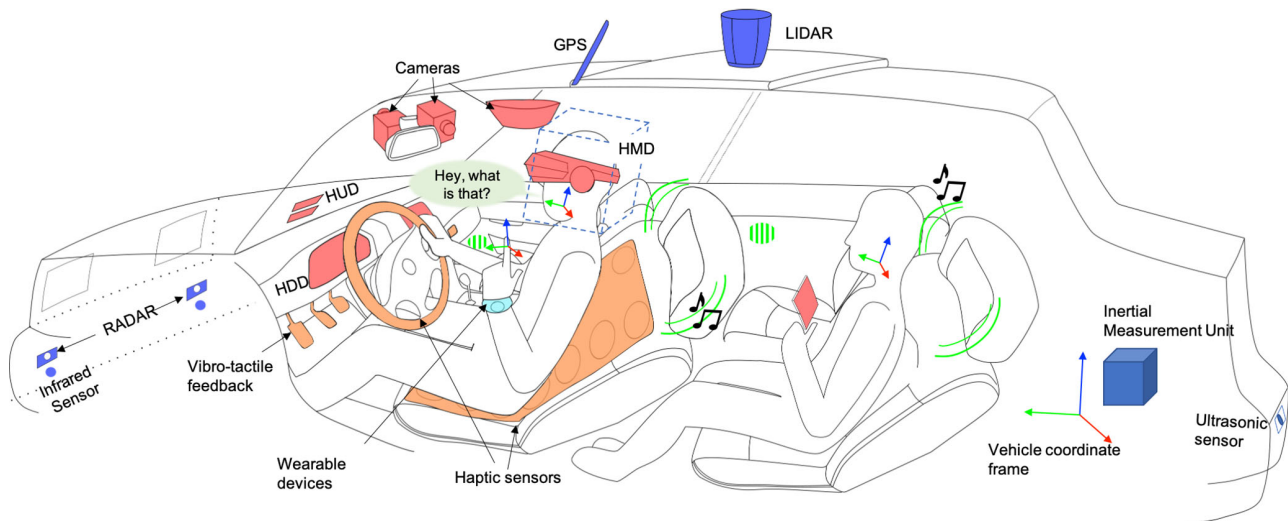


Figure 1. Schematic of in-vehicle interaction. Visual interfaces and sensors are shown in red, haptic interfaces in orange, exterior sensors in blue, and audio interfaces in green. Head pose recognition shown in dashed blue lines as well as various modes of interaction such as eye gaze, hand gestures, and speech are displayed.

world such as vehicle-to-pedestrian (V2P),^[11,12] vehicle-to-vehicle (V2V),^[13] vehicle-to-infrastructure (V2I),^[14] and vehicle-to-everything communication (V2X).^[14–16] These reviews present the challenge arising due to introduction of AVs into the traffic scenario, for example, interaction between pedestrians, cyclists, and other vehicles on the road. With comprehensive analysis of various technologies and strategies for in-vehicle interaction, this article complements the aforementioned reviews on the topics closely related to the emerging concept of HVI. In this review article, we focus on intelligent interior interaction between a vehicle and the driver or passengers. We review the technologies and interfaces for interaction, in particular haptic interfaces which are growing in importance in addition to visual and auditory interfaces.

Furthermore, we surveyed the state-of-the-art methods and techniques used for sensation, perception, and interaction. Such an analysis is timely considering the growth in the field. This is evident from **Figure 2**, which shows exponential growth of research articles published on AVs (Source: Web of Science. Keywords: “autonomous vehicle” or “autonomous driving”) during 2000–2020. Although the number of articles on HVI (Source: Web of Science. Keywords: “human–vehicle interaction” or “driver–vehicle interaction” or “interactive vehicle,” 9490 articles) published during the same period is much lower, their steady increase during the same period reflects the growing shift toward HVI and therefore the growing importance of technologies and methodologies for in-vehicle interaction.

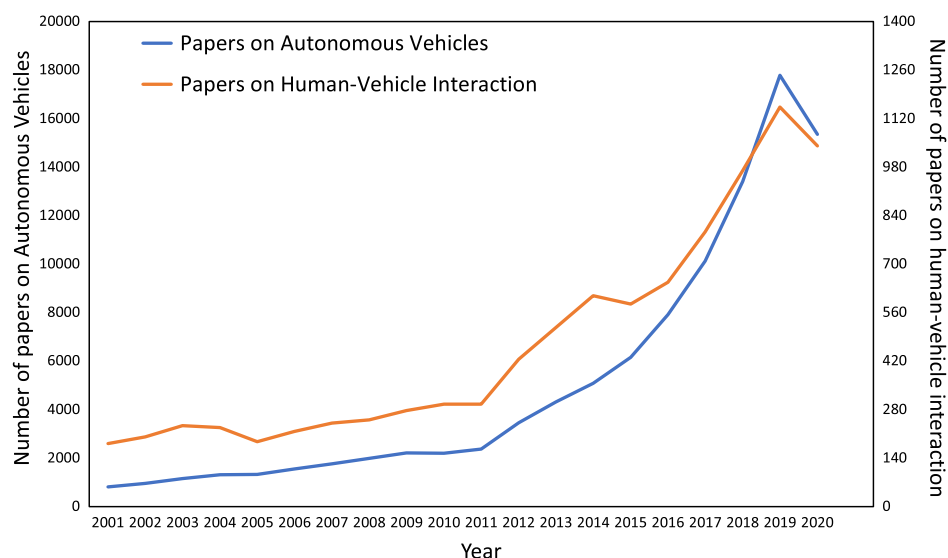


Figure 2. Comparison between number of papers on AVs and human–vehicle interaction over the last 20 years. Source: Web of Science. Keywords: “autonomous vehicle” or “autonomous driving”; “human–vehicle interaction” or “driver–vehicle interaction” or “interactive vehicle.”.

The organization of the article is as follows: relevance and types of in-vehicle interaction are described in Section 2. Various sensing technologies for HVI are described in Section 3. An in-depth discussion about the methodologies that use these sensing technologies to enable various interaction modes is described in Section 4. Section 5 presents the challenges for natural and intuitive interaction as well as attempts to provide future directions of research in this field. Finally, Section 6 provides the conclusions. In the Appendix, we provide a list of acronyms used in this article to aid the reader. We hope this article can provoke engineers, designers, and researchers toward developing novel multimodal interfaces and intelligent methods for interaction.

2. Level of Automation and In-Vehicle Interaction

There are six levels of automated driving systems ranging from level 0 (completely manual) to level 5 (FSD) automated systems. These are shown in **Table 1**. At each of these levels, there are distinct problems for HVI and roles of the driver, which can be classified as follows^[17]: 1) primary tasks involve maneuvering the vehicle such as steering and braking; 2) secondary tasks involve maintaining safety functions such as operating turn signals and windshield wipers; and 3) tertiary tasks involve all other comforts and operating in-vehicle infotainment systems (IVIS).

In level 0–3, tertiary tasks which involve driver–vehicle interaction should not cause distraction from the primary and secondary tasks. The main task of the driver is to comprehensively analyze the information of vehicle’s movement state and traffic condition and make the correct driving strategy and make corresponding driving action. For instance, the SAE established the so-called 15 s rule, i.e., any tertiary task that takes more than 15 s to conduct while stationary is not allowed while the vehicle is in motion.^[18] However, with the advent of level 4 and level 5, the primary tasks for the drivers and passengers are to perform nondriving-related tasks such as in-vehicle infotainment, working, and so on and secondary tasks will be supervising the AV and traffic situation.^[19] Therefore, depending on the driving context the vehicle needs to interact with the humans in certain ways according to established regulations and personal preferences of the drivers. Furthermore, the driver–vehicle interface is also responsible to smoothly handover from full self-driving to

manual control when necessary and regulate other interaction modalities upon temporal priority. The role of intelligent vehicle assistants (IVA) or intelligent driver assistants can vary dependent on the current task, the level of automation, and the cognitive state of the driver. There is no clear demarcation of the role to undertake and the IVA can perform a variety of roles in unison depending on the needs for the user. Some of the major roles and objectives of the IVA can be characterized as follow.

2.1. Entertainment

Studies have shown that to increase the acceptance of AVs, the AV must engage the users in activities to avoid boredom while removing the cognitive load and allowing them to relax.^[20] From large screen displays for working^[21] or watching movies to immersive technologies such as augmented reality (AR) and virtual reality (VR) for gaming,^[22] infotainment, and so on, IVA can enrich the joy of ride for the passengers.^[23] The objective is to transform the vehicle into a “living space” for mobile working, gaming, socializing, relaxing, and even to sleep.

2.2. Safety

One of the major objectives of autonomous driving is the increased safety for the driver, passengers, and other road users. In-vehicle interfaces play a crucial role in assisting the driver to enhance the level of awareness and safety. In level 5, an AV can handle all possible situations and the driver need not pay attention to the road situation. Up to level 4 automation, the driver will have to take control of the vehicle in certain situations. Currently, level 2 vehicles are entering the automotive markets. Therefore, driving assistance systems are provided through in-vehicle interfaces which are capable of recognizing implicit behaviors such as driver distraction,^[4] drowsiness detection,^[6] emotion detection,^[5] and various other activities of the driver. Different modalities (visual, auditory, haptic, or multimodal sensing) may be used to recognize the driver’s behavior and activities and make inferences on the situational awareness of the driver. The driving assistance systems can even provide additional features such as blind-spot assistance, parking assistance, navigation assistance, collision detection, and so on. A particular problem of level 3 driving is the take-over requests (TOR), wherein the AV signals through various in-vehicle interfaces to the driver to take over

Table 1. SAE J3016 levels of autonomous driving.^[1]

SAE automation level	Name	Description
0	No automation	Driver has full control and responsibility for all aspects of vehicle control even if automated system may provide warnings.
1	Driver assistance	Driver has full control of vehicle; the automated system provides features such as lane keeping assistance or adaptive cruise control.
2	Partial automation	The automated system can take control of vehicle functions such as accelerating, braking, and steering. Driver is required to continuously monitor the automated system for possible intervention.
3	Conditional automation	The automated system takes over vehicle function as level 2. The driver is necessary to be present, however not required to monitor the automated system at all times.
4	High automation	The automated system is capable of performing all driving functions under certain conditions. The driver may have the option to control the vehicle.
5	Full automation	The automated system is capable of full self-driving in all possible conditions. Driver intervention is not required.

control of the vehicle. The intelligent vehicle assistant is responsible for identifying the best modality (visual, haptic, or auditory) to notify TOR to the driver. In very time-critical situations such as impending collisions, a highly AV might require the driver to disengage to autonomous maneuver the vehicle to safety. Again, the selection of correct modality for communicating to the driver is important. Recognition of driver's mental state is key enabler for increasing safety during assisted driving.^[9]

2.3. Personalization and Recommendation System

By tailoring in-vehicle services to user's preferences, learnt over the long term, it may be possible to enhance user acceptance and usability of AVs. Some vehicle OEMs such as Land Rover have proposed the concept of "self-learning car"^[24] that leverages artificial intelligence (AI) to learn user habits. The IVA should be capable of identifying the distinct characteristics of every individual and personalize the behavior accordingly, such as setting the in-vehicle climate control, providing personalized music for each passenger without disturbing the others and seat configuration, and so on. Upon learning the habits and preferences, the recommendation system can provide precise recommendations to the users.^[25] The IVA can also assist in safety, for instance, recommending to take a break during long drives when drowsiness of driver is detected.

2.4. Natural and Empathetic Interaction

For a natural and intuitive interaction with a human, the IVAs needs to understand the subtleties of human communication such as facial expressions, eye gaze, hand gestures, voice tonalities, contextual cues, body language, and so on. To ensure such a natural interaction, multimodal sensing is needed along with perception algorithms capable of handling the different kinds of multimodal data. Furthermore, active sensation and perception are necessary so that unnecessary data collection is avoided, and framework can select the most informative data for the task being modeled. Contextual decision-making is necessary to adapt the vehicle behavior or in-vehicle interfaces.^[3] In additional, empathetic interaction entails the IVA to keenly observe the emotional state of the user such as anger, happy, sad, and so on and provide recommendations or alter the interactions to adapt to the user.^[5] Studies have shown that such empathetic interactions can increase trust, usability, and acceptance of IVAs.^[25]

The rich interaction with the intelligent vehicle is underpinned by various sensing technologies discussed in

Section 3. These sensing technologies are typically coupled with basic human sensory modalities for interaction such as visual, auditory, and haptics.

3. Technologies for Interaction

To learn about its environment and the passengers, the AVs require accurate sensing technologies. The AVs need to recognize human actions and intentions and provide interfaces for feedback. Typically, the input (humans to AV) and output (AV to humans) modalities are dependent on the human sensory capabilities. Humans can interact with AVs explicitly by performing specific actions to recognize human's intention. In this section, we review the technologies that enable intuitive interactions between humans and vehicles of the future.

3.1. Vision-Based Technologies

Visual feedback is one of the primary senses for humans. The visual touchscreen displays are now common in modern vehicles and provide information related to the vehicle, navigation, infotainment, and control functionalities such as in-vehicle climate control, parking assistance, sunroof control, and so on.^[26] The touchscreen displays have rapidly replaced the buttons and knobs in the head unit. Different types of displays that are being used or explored for current and future vehicles are shown in **Figure 3**. Displays offer multiple functionalities that can be used alongside third-party applications as well just as with smart phones. They can provide over-the-air updates to fix bug reports and can be customized to user requirements. As an example, Mercedes-Benz has recently showcased its concept car termed the VISION EQS featuring a central display seamless emerging from the central console along with separate side display units for the passengers in front and back seat allowing personalized viewing for each passenger.^[27] Unlike conventional buttons and knobs, these head-down displays (HDDs) shown in Figure 3a require visual attention causing distractions. Placed in the central console, the HDDs require the driver to take eyes-off-the-road to view information from the display and cause increased cognitive load while performing a primary task.^[28]

The alternative to this is the head-up displays (HUDs), which are designed to project information through the windshield onto the road ahead described through Figure 3b. Traditionally used in aircraft to project information that would be seen on the instrument panel, the HUDs have permeated the automotive

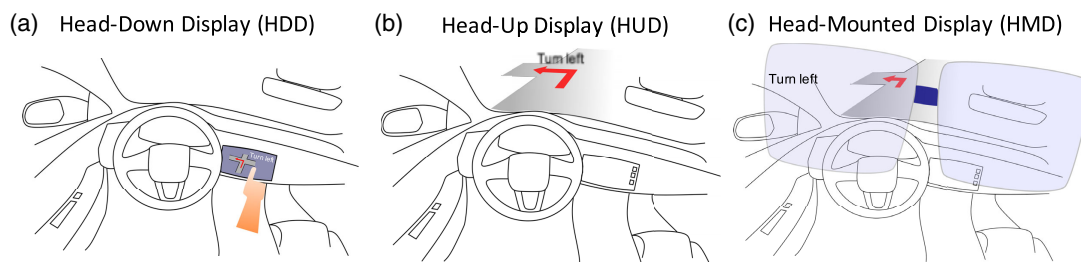


Figure 3. Different types of display ranging from a) touchscreen-based head-down displays (HDD), b) head-up displays on the windshield using AR technology (HUD), and c) HMDs.

industry, and have become a common feature on a long list of new cars. By projecting information onto the road, where the driver is already looking, the HUD systems limit the amount of distractions for a driver and therefore improve safety. The HUDs often complement the conventional dashboard displays and provide supplementary driving-related information. For example, the information that they project could be anything from which radio station is playing to your speed limit or they can assist the drivers to navigate effectively in accident-prone and low visibility situations.^[28] The HUD technology is quickly being refined because it is being seen as the next major advancement in interior automotive technology. For example, the use of AR holographic technology for HUD has recently attracted interest of automotive industry. As such, the AR-HUD technology is not new as it has been used for Boeing 727 class commercial transport in 1990s.^[29] The AR-HUD uses optical projection to present virtual information enhancements in the driver's line-of-sight without frequent adjustment between the real world and dashboard/navigation data.^[30] Automotive manufacturers have used the windshield of the car as a holographic display unit to provide road intersection guidance,^[31] ego-lane analysis,^[32] and virtually see through objects that are obstacles in the path.^[33] The AR-HUD technology could also find its way to infotainment in future AVs. For example, the novel touch enabled holographic display based on the frustrated total internal reflection (FTIR) principle could be fitted in the car to allow interaction with midair holographic objects through the sense of touch.^[34,35] The growing applications scope of such smart displays will also lead to continuous innovation in technologies and materials used. For example, high-performance materials such as polymer films, inks, and adhesives may emerge as the key components that will transform in-vehicle panels, touchscreens, and many other applications featuring smart displays.^[36,37] Likewise, such applications are drivers for research in transparent electronics or flexible electronics for roll-up displays.^[38–41]

The head-mounted displays (HMDs) is another class of displays gaining popularity among vehicle manufacturers (Figure 3c). HMD is a display device worn on the head or as helmet that has a display optic device in front of one (monocular) or both eyes (binocular). A typical HMD has one or two small displays, with lenses and semitransparent mirrors embedded in eyeglasses (also termed data glasses), a visor, or a helmet. AR goggles or smart glasses worn by the drivers can offer a plethora of use-cases such as providing vehicle speed and traffic information, navigation assistance in poor visibility conditions, parking assistance coupled with rearview cameras, digital assistants that can provide information such points of interest, messages, and so on during highly automated driving (HAD). Furthermore, AR displays can provide “see-through” technology wherein image processing techniques are used from various cameras placed around the car to superimpose digital images of unseen objects over the actual scene visible to the driver.^[42]

Autostereoscopic display (S3D) is another alternative, which allows the perception of 3D images without the need for any special headgear. Such 3D displays are composed of a color imaging liquid crystal display (LCD) and a monochromatic barrier liquid crystal display optically bonded. Vertical transparent and light blocking stripes will alternate on the barrier LCD. The subtly opposite viewing positions of the right and left eyes toward

the screen create two interwoven viewing areas on the imaging LCD. This generates the feeling of depth in the eyes of the viewer.^[43] S3D displays have been used in cars such as Mercedes-Benz User Experience (MBUX)^[27] which features autostereoscopic displays.

Visual sensors including monocular and stereo RGB cameras, depth cameras (RGB-D), time-of-flight laser sensors, and so on are increasingly present in the vehicles these days. While exterior perception for highly automated driving may include sensors such as RGB cameras, light detection and ranging (LIDAR), radio detection and ranging (RADAR), and ultrasonic sensors, they can also be used within the car for interior perception such as gesture recognition,^[44] human activity recognition,^[45] head pose estimation,^[46] driver distraction or fatigue detection,^[47] and so on. The interior cameras are located near the back-view mirror and on the vehicle's rear flank. To capture a 180°–360° view of the surroundings, more than one camera or pan-tilt-zoom (PTZ) cameras are needed in some applications. As RGB cameras are affected by illumination, infrared (IR) cameras are also used for in-vehicle applications such as head pose estimation^[48] as they are robust to lighting conditions. RADAR, in particular impulse radio ultra-wide band (IR-UWB) radar, has been used in-vehicle detection of babies or pets to prevent accidental locking inside the car.^[49] Exterior sensors such as LIDAR, RADAR, and stereo-cameras are used for pedestrian detection and intention estimation. Multiple cameras and LIDAR on the vehicle need to be calibrated with respect to the vehicle coordinate frame.^[50] As the positions are fixed relative to the vehicle, calibration may be done when necessary and typically use special calibration targets such as boxes,^[51] planar checkerboards,^[52] and so on. Recent works have also explored laser sensor calibration in a target-agnostic approach in the robotic domain which can be extended to the automotive domain.^[53] A comparison between various exteroceptive sensors including visual sensors is shown in Table 2.

3.2. Haptics or Touch-Based Technologies

There has been a clear trend in the automotive industry for replacing buttons and dials with a central touchscreen display providing multiple functionalities. As discussed in previous section, the touch-based visual displays have become the new norm. The touch-based control in visual displays is generally achieved through transduction mechanisms such as resistive, capacitive, and surface acoustic waves.^[35,54] Among these the capacitive sensing is most popular in touchscreens due to simple electronics and multitouch sensing.^[55] However, considering that the visual display may distract drivers, researchers have started to explore other intuitive and nondistractive methods for providing feedback to the drivers. In this regard, tactile or haptic interfaces offer several advantages such as) nonvisual: secondary information can be communicated to drivers and tasks completed without performing glances toward any screen and taking eyes off the road; 2) natural control: controls using touch are known to be more intuitive and natural to humans and require less cognitive load; 3) privacy: communication can be discretely performed between the person and the car without the need to be displayed or announced; 4) spatial resolution: touch sensing in humans have a high spatial resolution and it has been shown that humans

Table 2. Comparison of exteroceptive sensors.

Sensor	Measured information	Depth range	Resolution	Sensitivity to illumination	Sensitivity to weather conditions	Sensitivity to transparency
Camera	RGB	–	High	Yes	Yes	Yes
Stereo-camera	RGB, depth	100 m	High	Yes	Yes	Yes
LIDAR	Depth	200 m	High	No	No	Yes
RADAR	Depth	<180 m	Low	No	No	No
Ultrasonic	Depth	10 m	Low	No	Moderate	No
Event camera	Events	–	–	No	Yes	Yes
IR camera	Short wave IR images	–	Medium	No	Yes	Yes
Thermal camera	Long wave IR images	–	Medium	No	Yes	Yes

can distinguish patterns up to 13 nm from a smooth surface^[56]; and 5) reaction time: it has also been shown that tactile feedback is faster than visual feedback.^[57]

The technical requirements for haptic or tactile interfaces within the car include large area distributed multimodal sensing, good spatial resolution, bendability and flexibility, fast response time, low hysteresis, high durability to repeated contacts, and robustness.^[58,59] The tactile sensors can be embedded within the vehicle interiors where there is usually contact with the humans such as steering wheel, dashboard, seat, head-rest, and so on. Dynamic force sensing, distributed pressure sensing, and point of contact localization are required for extracting tactile information.^[60] Capacitive sensing is a popular choice due to high sensitivity, simple read out electronics, robustness, and low power consumption.^[54] Novel sensors which are extremely thin, flexible, and stretchable can be designed by printing conductive material into thin stretchable sheets.^[61] These sensors can be integrated into vehicle interiors that have complex shapes and structures. A detailed review on tactile sensing technology particularly in the robotic application domain, but also suitable for automotive applications, can be found in previous studies.^[58,62,63] Similar to capacitive touch sensing, proximity sensing can detect conductive objects, such as humans at close proximity. Therefore, it can be used to detect in-air gestures as well as touch-based gestures.^[64] A capacitive sensing layer can be embedded under a nonconductive material such as leather, plastic, or wood and these materials are usually present within a car. For example, an active sensing armrest, reported using capacitive proximity sensing, combines limb detection and recognition of multiple gestures.^[65] Along the same lines, a comprehensive review of the usage of capacitive proximity sensing in smart environments is given in Braun et al.^[66] Other alternative for haptic feedback is the vibrotactile inputs, which can be provided to alert a drowsy or distracted driver.^[67,68] For instance, by providing tactile alerts to the driver through sensorized seat, TOR in the context of highly automated driving (HAD) have been investigated.^[69] Six tactile sensors from Engineering Acoustics Inc. were placed on the seat such that directional tactile signals will be provided to the user on either side of the thighs.

The novel device capable of providing tactile sensing as well as vibrotactile feedback integrated into one single module is another interesting example.^[70] The vibrotactile component of the device can provide feedback at frequencies (10–200 Hz) that are within the perceivable tactile frequency thresholds of the

human hand. Such devices could be possibly integrated into areas of the car such as steering wheel and armrest wherein sensing and feedback is necessary. Pressure sensing can also be combined with other modalities such as temperature. For example, a multifunctional touch sensing electronic skin with stacks of capacitive pressure sensors and temperature sensors has been reported.^[71] This flexible skin is capable of sensing pressure >10 kPa and temperature up to 80 °C with fast response (2.5 s) and recovery (4.8 s) time. Human activity monitoring such as driver posture monitoring use-cases require the calculation of mechanical strain and deformation. In this regard, the highly sensitive flexible strain sensor using conductive polymer polystyrene sulfonate (PEDOT:PSS) microchannel inside a polydimethylsiloxane substrate is worth noting.^[72] Such a device could also be used as wearable sensors for the driver to measure muscle fatigue. Haptic feedback is also provided using large tactile displays that are present in modern-day vehicles. As an example, the Active Sensing Technology system designed by immersion using piezoactuators can provide precise and high-fidelity haptic feedback on display units.^[73] Integrating haptic feedback over large area surfaces is not trivial as it requires precise calculation of mass, direction, and localization of the haptic effects. Furthermore, haptic feedback can also be provided in accelerator pedals^[74] and steering wheels in order to minimize driver distraction.^[75–77]

Smart interactive surfaces present another new direction providing seamless integration of user interfaces into the interior surfaces of vehicles, which were originally designed for purely aesthetic purposes only. In the past, there was a clear demarcation between the interior surfaces which are meant for decorative purposes and those that provide control inputs or feedback outputs. However, this clear boundary has blurred recently with the advent of smart interactive surfaces which have control elements integrated into decorative surfaces. A combined use of decorative lighting elements, capacitive switching technologies, and tactile actuation integrated with force sensors embedded under a textile material for a functionalized door component has been demonstrated in Blomeyer and Schulte-Gehrmann.^[78] In this case, the capacitive switching functions for seat memory and seat heating are visible only when search lighting is activated. Furthermore, tactile feedback to activate the switch is supplemented by visual feedback through colors of different switching units. Such technologies can be used only when necessary and exhibit the functionality of secret-till-lit.^[78] Other example of smart soft surfaces,

with control elements seamlessly integrated into textile for seats and other soft interiors, has been demonstrated by Yangfeng automotive interiors.^[79] Recently, BMW has also demonstrated the intelligence surface called “Shy Tech,” for both interior and exterior surfaces.^[80] This intelligence surface contains cameras, RADAR, and many sensors, takes on digital functions, and has a self-healing effect.

Smart textiles as sensing elements have also been explored to develop intelligent interior surfaces as textile is typically used in the car for seat covers, seat belt, roof top, door panel, and parts of the dashboard. The fiber-based capacitive and resistive sensors are quite common. Among capacitive textile sensors, the common approach is to include single element sensors, conductive fabric stripes, e-broidery, printed patterns, coated fiber, and hollow structured fiber.^[81] Single element sensors can sense contact at one point. An example includes capacitive sensors for seat, wherein electrodes built with conductive textiles were arranged on both sides of a compressible spacer, forming a variable capacitor.^[82] The hollow structured fiber often uses the air inside the fiber as a spacer and can be very robust to repeated sitting and washing typical of car seats. Conductive fabric stripes, knitted or woven by conductive yarn, can also be used as both support structure and conducting electrode. For instance, silver yarns-based sensor knitted into a fabric has been demonstrated with low hysteresis,^[83] as shown in **Figure 4c**. The patterns of conductive ink printed on the fabric to form small sensing elements, which can be scaled at low cost, are another attractive route for smart interior. The printed patterns are also robust to washing. An example for this approach is the printed textile-based strain sensor, which has been demonstrated to measure finger angle and the movements of the pharynx when speaking, coughing, and swallowing.^[84] Another variant of smart textile sensors is based on coated fibers, wherein the fabric is knitted or woven by specially coated fiber and the capacitive change between two fibers gives a measure of pressure variation.^[85] Xsensor^[86] and Tekscan^[87] are some of the manufacturers of textile-based pressure sensors which can be used with vehicle seats for monitoring driver ergonomics, as shown in **Figure 4a,b**.

On the contrary, the resistive textile sensors often include single element sensor made of various screen-printed transducer materials, e-broidery, or nanofiber.^[81] Examples include piezoresistive textile sensor wherein the transducer material is sandwiched between a nonconductive textile and a conductive

grid-structure.^[88] The sensor can detect multitouch points or can be multimodal. As an example, the stretchable and weavable piezoresistive multimodal textile sensor capable for distributed pressure and strain sensor have been reported recently,^[89] as shown in **Figure 4d**. Such multimodal sensors can be advanced to have wireless data transmission by using innovative methods such as using stretchable antenna as strain sensors.^[90] In fact, there are other interesting avenues such as self-powered sensors that have been explored in other application areas. For example, triboelectric nanogenerators have been demonstrated to show varying power output with applied load, indicating that they can be used as both energy harvesters and pressure sensors.^[91,92] Considering the increasing demand for energy, such solutions can add significant value to the automotive sector. The conductive cotton and jute fiber- and resistance-based sensors have also been explored for temperature and humidity sensing and energy storage devices.^[93] Such sensors, made from sustainable materials, offer good solutions for interiors such as foot mats where fibers are relatively thicker. With fiber-based supercapacitor foot mats, one could implement interesting, distributed energy paradigm.^[94] In contrast to using centralized batteries in future electric vehicles, the distributed energy can help to generate more space inside the vehicle and improve the passenger comfort.

Other examples of in-vehicle interaction include midair gestures while having realistic haptic feedback without touching any surface. These technologies can create rich, 3D shapes and textures that can be felt. To this end, aerohaptic and ultrasonic wave-based methods have been explored.^[35] In this regard, the ultrasonic wave-based IVIS by Ultraleap and Harman^[95] is noteworthy. This system responds to the driver's gesture commands with tactile sensations confirming instructions.

3.3. Auditory Technologies

Auditory modality plays a central role in environment perception for driving-related tasks as, for some situations, no other modality can replace the hearing modality such as honking or the sirens from an oncoming emergency vehicle. Typically, the auditory modes of interactions include in-vehicle microphones to receive human commands in the context of voice-user interface and loudspeakers for providing in-vehicle infotainment and feedback suggestion in the case of voice-based interaction. Auditory perception can complement traditional visual and haptic-based

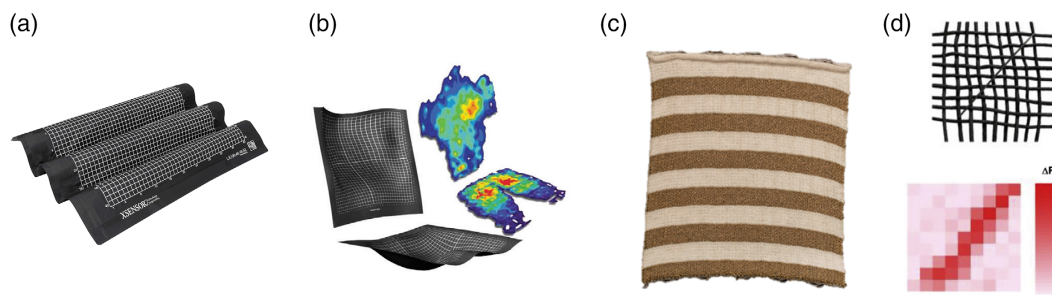


Figure 4. Examples of commercial and research prototypes of textile-based sensors. a) Xsensor. Reproduced with permission.^[86] Copyright 2021, Xsensor Technology Corporation. b) Tekscan. Reproduced with permission.^[87] Copyright 2021, Tekscan, Inc. c) Knitted capacitive textile sensor. Reproduced with permission.^[83] Copyright 2019, Wolters Kluwer Medknow Publications. d) Piezoresistive carbon nanotube-based weavable and stretchable sensor. Reproduced with permission.^[89] Copyright 2020, Elsevier Ltd.

interaction to perceive and interact with humans and require research attention. Novel technologies for auditory perception are discussed in this section.

Considering in-cabin infotainment and audio, passengers may want to listen to an audio of personal preference with minimum cross interference from the audio of the person seated next to them. Although it can be achieved through headphones, greater comfort and joy of ride can be achieved by creating personal audio zones through loudspeakers around the headrest or in proximity to the passenger.^[96,97] The idea of personal audio zones has been explored by many automotive manufacturers such as Hyundai Motor Group.^[98] Other examples of such technologies include loudspeaker array on the ceiling of the car cabin to generate independent listening zones in the front and rear seats at higher frequencies.^[99]

The microphones typically present within the car can also be placed on the exterior to detect sirens, vehicles in proximity, and even pedestrians.^[100–102] Other interesting use-cases of acoustic sensors include learning algorithm for audio-only odometry that only measured the acoustic signals from external microphones with good prediction accuracy.^[103] This system was not affected by the scene's appearance, lighting conditions, and structure. The experimental evaluation demonstrated significant resilience to environmental noise and it can be used as an auxiliary modality for visual mode for egomotion estimation.

3.4. Physiological Sensing Technologies

Future AVs are expected to provide rich experience that goes beyond infotainment. In this regard, real-time health status and mental state make important set of areas considering safety and well-being of the passengers. A wide range of physiological sensors developed for wearable applications are being repurposed to measure specific health-related data and reveal the cognitive state so that the distraction levels of the drivers can be reduced. These sensors measure physiological parameters such as heartbeat, blood pressure, muscles movements, eye-tracking, and so on. Detailed description of such sensors, particularly for wearable systems, is given elsewhere^[36,59,104] and some of these commonly used in automotive context are described later.

1) Electroencephalography (EEG) is a monitoring method to record electrical activity on the scalp that has been shown to represent the macroscopic activity of the surface layer of the brain underneath. Research in driver distraction has shown that EEG signal is one of the most conclusive measures of measuring driver fatigue or sleepiness.^[105,106] EEG is characterized by four activity bands depending on the frequency range as beta, alpha, delta, and theta. The onset of sleepiness is characterized by the Theta waves and sleep state characterized by Delta activity.^[107]

2) Electrooculography (EOG) provides information about both eye movement and blink patterns. Typically, cognitive alertness in drivers is characterized by rapid eye movements and onset of drowsiness causes slower movements and longer blink rates.^[108] For instance, eye movement monitored using EOG can provide an accuracy of 80% for driver drowsiness detection.^[109]

3) Electromyography (EMG) is a technique used for recording electrical activity in muscles. For instance, in Katsis et al.^[110] the authors observed a marked decrease in amplitude and frequency

on EMG signals with the onset of driver distraction and therefore can be used as an effective method to measure alertness levels.

4) Electrocardiography (ECG) monitors the cardiac activity and the heart rate. ECG is easier to capture and can provide a variety of signals that can reveal the alertness state of the driver.^[111] Furthermore, heart rate can also reveal the mood of the driver as excited or angry (high heart rate) and calm (normal) state.

5) Body temperature sensing can reveal the comfort levels of the drivers and passengers and can help regulate the internal temperature of the car as per the personal preference of the driver and passengers. A wide variety of wearable systems have been reported for measurement of body temperatures^[112–114] and some of the advanced solutions can wirelessly transmit the data to smartphone or to electronic units in the vehicle.

3.5. Proprioceptive Sensing Technologies

Another important category is the proprioceptive sensing which measures the internal state of the vehicle.^[115] For example, vehicle states such as speed, acceleration, and yaw must be continuously measured for safe operation. Commonly used proprioceptive sensors include inertial measurement units (IMU) for determining the vehicles acceleration, heading sensors (gyroscopes and inclinometer), wheel encoders for counting rotations of the wheel, altimeters to measure altitude, and tachometers for calculating the revolutions per minute of the shaft. Several such sensors are described in detail elsewhere.^[115]

4. Methods for In-Vehicle Interaction

In-vehicle interactions in the next-generation vehicles promise to promote driver situation awareness, trust, comfort, better user experiences, as well as usability and safety. Traditional in-vehicle displays are expected to be expanded beyond graphical user interface (GUI) displays with multimodal interfaces, including auditory,^[116] tactile/haptic,^[117] gesture,^[118] wearable sensors,^[119] and AR/VR/mixed reality (MR) technologies.^[120] Vehicles are being equipped with multimodal sensing technologies, as described in Section 3, to ensure accurate predictions for in-vehicle interactions. Furthermore, driver or passenger monitoring is crucial for interaction. In-vehicle interaction systems need to estimate and infer driver/user actions, states such as fatigue or drowsiness, cognitive state of the driver, and emotions of the users. Typically, users can interact with an intelligent vehicle implicitly as well as explicitly. In the implicit communication, users or drivers conduct their behavior in its own right and an observer can infer the state or intention of the user performing certain behaviors.^[121] Examples include driver fatigue recognition, emotion recognition, and even posture or pose estimation which can convey certain cues to the intelligent vehicle. On the contrary, examples of explicit communication where the users intend to communicate with the vehicle or vice versa include voice commands, gestures, and communication through haptic and display interfaces. At times, the implicit cues can also assist while interacting explicitly with the humans. In this section, we review the state-of-art methods used for in-vehicle or interior interactions.

4.1. Implicit Interaction

Recognizing driver and passenger behaviors and activities has far reaching implications for in-vehicle interaction systems and safety functions. The human-machine interaction can be guided to the most appropriate modality (visual/audio/haptic) at each moment if the driver's intentions are correctly classified. The safety functions, such as airbags, steering, brake, and crash avoidance patterns, can be tailored to the best in-time deployment if the car knows the full body position (sitting, lying, etc.) of its passengers. During semiautonomous mode, the AV will use in-vehicle movement and driver attention/disengagement in the driving task to determine if a handover can be achieved safely or whether a safe stop should be performed instead. The current best modality for in-vehicle warnings can also be optimized based on the situation in the cabin (e.g., not rely on visual HMI when driver is reading or looking at a phone). From the interaction perspective, the ride in an autonomous or semi-AV can be adapted to the state and activity of the driver and passengers. Furthermore, with tracking of face expressions, gestures, and body position, the emotional state and response of the driver and/or passengers can be used to evaluate the automated vehicle's actions in traffic. Research on safe human-robot interaction and in particular human activity recognition techniques is particularly useful even in the automotive context.^[122] Mapping of all passengers in the AV will enable new methods of understanding how not only social interaction between passengers, but also between passengers and the intelligent car will look like in the future.^[123] Upon detecting driver distraction or fatigue, the intelligent vehicle assistant (IVA) may choose to provide visual or vibrotactile alerts. Furthermore, the IVA can also engage in conversation with the driver to keep them alert. If a nonoptimal body posture is detected, the actuated interiors can nudge the driver to a correct body posture for optimal driving attention. Similarly, if the driver's mood is found to be angry or sad, the IVA can recommend soothing music, control the in-vehicle temperature, and create a relaxing environment. Furthermore, if the driver is found to be incapable of controlling the vehicle, the AV can disengage the driver to take the driver to safety as well as communicate to other vehicles nearby (V2V) and related authorities (V2X).

4.1.1. Driver Fatigue and Distraction Recognition

Driver distraction is one of the major causes of accidents on the road and the US National Highway Traffic Safety Administration (NHTSA) estimate up to 25% of the road accidents happening due to some form of driver distraction.^[124] Driver distraction or inattention can be defined as "Driver inattention represents diminished attention to activities that are critical for safe driving in the absence of a competing activity."^[125] Distraction and fatigue are two common forms of driver inattention leading to driving accidents.

Prior works have identified various distinctions in the types of distractions such as performing secondary or tertiary tasks with hands (manual), eyes (visual), and/or mind (cognitive) off-the-road.^[126–128] While these being the major sources of distractions, there can be also other inputs such as auditory stimuli which

cause distraction. Furthermore, some activities can be a combination of the different types of distractions such as texting, while driving is a combination of manual, visual, and cognitive distraction. To detect driver distractions, various sensing technologies have been discussed in Section 3. We discuss the relevant methods and algorithms employed using the different sensing technologies and, furthermore, discuss the methods of fusing one or more types of sensing information in this section.

Exterior sensors such as vehicle IMU, GPS, and exterior cameras can be used to infer driving behaviors and patterns of erratic driving which can mean distraction. Vehicle speed is one measure that is crucial and must be used with other sensor measurements while inferring distractions. Deep neural networks (DNNs) with architectures such as YOLOv4^[129] and Faster R-CNN^[130] have been used with exterior cameras for detecting other vehicles and pedestrians in traffic as well as road signs. Identifying lanes and lane-keeping error can be a metric for detecting driver distraction.^[131] For instance, a logistic regression model has been used to distinguish between distraction driving and normal driving using lane-keeping errors.^[132] Steering wheel sensors also provide an indirect indication of driver distraction. A method to predict steering angle through a second-order Taylor series and the observed angles to calculate steering error has been demonstrated.^[133] The error increases when the driver is distracted.

In comparison, interior visual sensors such as stereo-cameras, IR cameras, ToF sensors, and RGB-D sensors have been widely used to recognize driver and passenger activities, intentions, and behaviors. State-of-the-art neural network architectures for detecting driver distraction such as VGG, AlexNet, GoogleNet, and ResNet have been compared and ResNet architecture seems to outperform other competing strategies.^[134] Even single image-based driver activity recognition to detect activities such as talking on phone, texting, eyes-off-road, rubbing eyes, and so on has been demonstrated using neural networks.^[47] An end-to-end network based on pretrained CNN VGG-19 architecture has been proposed which is robust to luminance, shadows, camera pose, and driver ethnicity to detect driver distraction.^[134] Furthermore, focused areas of the face can also be used to detect distraction, such as yawning can mean fatigue, eyes-off-the-road with eye-tracking, blinking rate, and so on. In contrast to DNNs, machine learning techniques such as support vector machine (SVM) have also been used to detect eye closures.^[135] To tackle the sparse labeled data problem, a novel framework termed Few-shot Adaptive GaZE Estimation (FAZE) capable of learning a compact person-specific latent representation of gaze, head pose, and appearance has been explored.^[136] Gaze estimation is then utilized to classify distracted driving and normal driving behaviors. The internal camera can also be used to detect hands-off-the-wheel activity. Other metrics to detect fatigue and distraction include percentage of eye closure (PERCLOS),^[135] eyes-off-road-time,^[137] and yawning and nodding^[135,138] using video sequences from RGB cameras. Furthermore, infrared cameras can be used to mitigate variation in lighting conditions.^[139]

Apart from visual sensing technologies, physiological wearable sensors such as electromyography (EMG), electroencephalogram (EEG), electrocardiogram (ECG), electrodermal activity (EDA), electrooculography (EOG), and heart rate sensors have also been used for detecting distraction. EEG signals can be

classified into four categories from vigilant to sleepiness while driving by training a SVM classifier.^[140] As EEG measurements require placing electrodes on the head of the driver, they serve very little practical use-cases; however, certain novel in-ear EEG may be deployed for real use-cases.^[141] ECG measures the electrical activity of the heart and the driver's emotions, mental activity, and body exertion affect the heart rate.^[142,143] EOG sensors are used to record eye movements and can be used to detect fatigue and sleepiness. For instance, a blink is detected when the contact between the eye's upper and lower lids lasts for about 200–400 ms, and a microsleep is detected if the eye remains closed for more than 500 ms and can be detected by the EOG.^[144] Studies have shown that on the onset of fatigue, the amplitude of EMG signals decreases gradually which can be detected by an EMG device.^[145–147] The physiological sensors are intrusive in nature which inhibits their practical usage. However, novel techniques such as the Neuralink (<https://neuralink.com/>) brain-machine interface wherein the wireless device is implanted directly into the skull may offer details of the brain activity and cognitive distraction which can assist driving as well as highly automated driving.

4.1.2. Emotion Recognition

Emotion recognition is critical for daily function in decision-making, communication, general mood, motivation, and even driving. Emotion recognition is a complex field of research requiring the use of physiological sensors and controlled studies thus increasing the complexity for in-vehicle driver emotion recognition. Recognition of emotions can also assist in detecting driver distraction and fatigue that has been discussed in Section 4.1.1, but can also elicit far more information about the driver behaviors and can aid in the objective of an intelligent vehicle that understands and responds to the driver.

Emotions can be classified into six categories as anger, disgust, happiness, sadness, surprise, and fear.^[148] Studies on driver emotion recognition have been conducted for recognizing the six main emotions or a smaller subset such as anger and happiness. With annotated datasets of head and face from RGB or IR cameras, supervised learning techniques can be used to classify emotions such as k-nearest neighbors (kNN),^[149,150] SVM,^[151,152] and DNNs.^[153,154] The studies provide frame-level predictions as well as window-level wherein different frame-level predictions are aggregated using a voting scheme. Using physiological signals instead, high dimensional signals are extracted using the variety of physiological sensors. Therefore, dimensionality reduction algorithms such as principal component analysis (PCA)^[155,156] and linear discriminant analysis (LDA)^[157,158] are used. PCA helps in finding a set of uncorrelated features that explain the variance in the original data while LDA fits the data with a linear combination of features while finding a linear function that discriminates classes. Upon dimensionality reduction, the popular methods used include SVM and Naive Bayes for supervised learning to classify the emotions. As biophysical data consist of time varying signals, recurrent neural networks which can capture the features from time varying signals have demonstrated good performance in emotion recognition.^[159] Speech signals are also used for emotion recognition as speech can

provide cues which can be used to decipher the emotion of the driver. Speech signal also consists of high dimensional signals that require dimensionality reduction techniques such as biophysical signals. Studies also reveal that features for speech signals such as pitch, energy, and intensity are useful for emotion recognition.^[160] The emotion state of the driver can also be inferred using the behaviors such as grip strength on the steering wheel is shown to vary when the driver was happy or angry.^[161] To identify the best framework for detecting frustration, five different supervised learning algorithms and their combinations (Bayesian neural network [BNN], SVMs, Gaussian mixture models [GMMs], multinomial regression [MNR], and GMM + SVM) have been compared. The results showed that SVMs and MNR performed the best for the task of frustration detection from a driver's sitting posture.^[162] A comprehensive review of emotion recognition in automotive use-cases is provided in recent reviews such as the work by Zepf et al.^[5] A summary of the techniques used for driver distraction, fatigue, and emotion detection is shown in **Figure 5**.

4.1.3. In-Vehicle Pose Recognition

For interaction, safety and analyzing driver or passenger activity, it is crucial to track body parts and the kinematics of the body such as head pose, body pose, position of hands, feet and trunk, presence of baby, and so on. Driver posture monitoring is also important from an ergonomic viewpoint of driving. Studies have shown that poor posture contributes to driver discomfort over long-haul drives.^[163,164] Therefore, pose recognition is analyzed separately in addition to the previously discussed sections on activity, emotion, and distraction detection.

Head pose estimation and tracking, as shown in **Figure 6a**, is critical for variety of use-cases such as driver monitoring, distraction detection, gaze detection, calibration of head-mounted devices, and so on. Template matching approaches have been demonstrated for head pose estimation and tracking.^[165] The head region is detected using feature extraction techniques on RGB images and then matched to an exemplar head image in training data. The pose of the exemplar image is the estimated head pose. Similarly, discrete head yaw and pitch values are detected by using a coarse-to-fine strategy using a quantized pose classifier in the work by Wu and Trivedi.^[166] Head pose can also be estimated by tracking key facial features such as eye corners, nose tip, and nose corners.^[146] The method demonstrates good performance when the face is in a full frontal view. As it is based on a monocular camera-based approach, performance degrades with single perspective, occlusions, or the presence of large movements. To tackle illumination issues of RGB cameras, IR camera-based approaches have been devised. For instance, a novel network termed Head Orientation Network (HON) and ResNetHG has been designed to estimate head pose from IR images in the AutoPOSE dataset.^[48] Monocular estimation is sensitive to rapid head movements, which can be tackled by placing multiple cameras in the field of view of the driver.^[167]

Structured-light depth cameras have also been proposed for driver pose estimation, as shown in **Figure 6c**. For instance, a graph-based algorithm to fit a 7-point skeletal model to the body of the driver using a sequence of depth frames has been proposed.^[168] Similarly, a time-of-flight depth sensor has been used

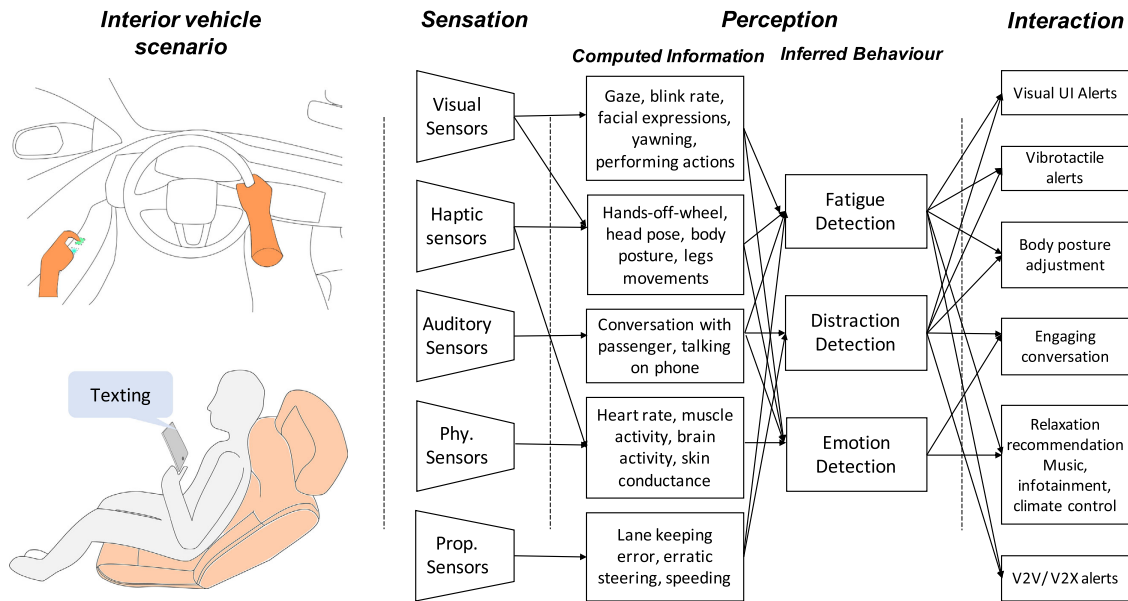


Figure 5. Framework for driver distraction, fatigue and emotion detection, and consequent interactions. Note that representative interactions are presented, as many more types of interactions may be implemented depending on the situation. Prop. Sensors, proprioceptive sensors; Phy. Sensors, physiological sensors.

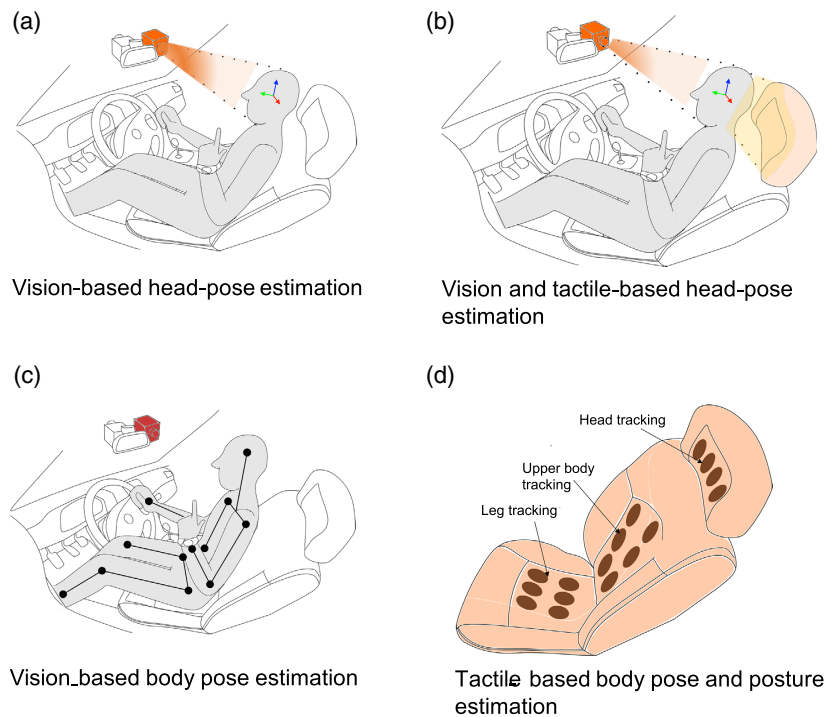


Figure 6. Various in-vehicle pose estimation use-cases demonstrated in the literature: a) head-pose estimation through visual sensors, b) vision and tactile-base sensing based head-pose estimation, full body-pose recognition and tracking through c) visual sensing and d) tactile sensors embedded on the seat.

to estimate the location and orientation of a driver's limbs, including arms, hands, head, and torso.^[169] Iterative closest point (ICP) algorithm was used for pose estimation using an

articulated 3D model of a human. However, as is well known ICP gets stuck in local minima and requires a good initial guess to execute the pose estimation process.

Apart from visual sensors, tactile sensors such as force-sensor arrays, tactile sensor skins, and proximity sensors have also been used for body pose and head pose estimation, as shown in Figure 6b,d. As discussed in detail in Section 3.2, tactile sensing arrays can be unobtrusively and seamlessly integrated into the seat. For instance, a seat with distributed pressure sensors on the backrest and cushions has been used to measure the driver sitting posture.^[170] Such force sensing arrays on seats have been used to measure driver and passenger body posture for safety critical applications such as airbag deployment.^[171] Similarly, pressure-maps fused with IMU measurements have been deployed to robustly recognize body pose.^[172] A sensorized headrest with capacitive proximity sensors has been used for head pose estimation and headrest servoing wherein the headrest can move according to the movement of the head of the user using a nonparametric neural network-based method.^[173] Other works demonstrate sensors such as ultrasonic sensors embedded in the headrest for head pose estimation.^[174] However to robustly estimate body and head pose, haptic sensors need to be combined with visual sensors in a multimodal fashion.

4.2. Explicit Interaction

In contrast to implicit interaction, users such as driver and passengers can explicitly interact with the intelligent vehicle assistants through engaging in conversation, proactively performing tasks, gesturing, and so on. The IVA also needs to consider various cultural and geographic aspects while proactively interacting with users. For instance, different gestures can mean different things in Italy or Japan. Furthermore, by analyzing the emotion of the user, the IVA can adapt its interaction mode suitably, thus leading to more empathetic interaction. Explicit interaction can take place via voice-based, display-based, haptic-based, and even multimodal interfaces as described in the sections hereafter.

4.2.1. Voice-Based Interaction

Voice–user interface (VUI) can drastically reduce driver distraction by allowing the driver to interact with the vehicle without taking the eyes-off-the-road or hands-off-the-wheel and reducing the visual cognitive load. Voice assistants are more commonly implemented for OEMs such as BMW Intelligent Personal Assistant with the “Hey, BMW!” prompt,^[175] Daimler’s MBUX voice assistant,^[176] and also integration of third-party assistants such as Amazon Alexa and Apple CarPlay. The digital voice assistant (VA) needs to understand the commands provided in a naturalistic way without relying on predefined key words that require prior training for the users. In-vehicle functions such as infotainment, climate control, communication such as making calls or sending texts, vehicle status (such as fuel left) and even providing assistance outside the car by syncing with other voice assistants.^[175] Navigation is one of the most used VUI tasks wherein drivers can request navigation directions on the go, without typing-in to a touchscreen console and looking at a display route in safety critical driving conditions.^[177]

During long drives especially during the night, it is known that engaging in conversation with fellow copassengers in the front seat can help keep the driver alert and awake.^[178] However, conversation increasing the cognitive load of the driver such as cell-phone conversations can be detrimental to driver alertness. Furthermore, passive listening may not be as effective as active participation in a conversation.^[178] Therefore, studies have been done wherein the driver and the VA are engaged in a casual conversation rather than information gathering sequence of voice commands and have shown that it can be effective in preventing driver distraction.^[179–181] Studies showed that short intermittent conversation with a VA can help increase driver alertness.^[179] Furthermore, driving simulator-based studies has been performed to show that VAs should provide assertive voices to grab driver attention.^[182] Guidelines for designing conversational VAs for engaging drivers in natural conversation are provided in the work by Large et al.^[183]

Voice assistants can also assist in alerting the driver for oncoming emergency vehicles. Sirens are a unique sound generated by emergency vehicles such as police cars, ambulances, and fire trucks. The siren sound is issued by the emergency vehicle to alert other vehicles and pedestrians to give way. Due to driver distractions or sound-proofing of modern cars, drivers may not aware of the oncoming emergency vehicle. Acoustic-based emergency vehicle detection method and providing acoustic and/or visual alerts for the driver is one method of tackling the problem. One of the initial real-time siren detection systems through a signal processing-based pitch detection algorithm was developed in the work by Meucci et al.^[100] The prior work on siren detection was extended to sound source localization and providing alerts to drivers with the proximity of the sound source.^[101] A CNN-based ensemble model was proposed to classify traffic soundscape to noise, siren sounds, and other vehicle sounds.^[102] The method demonstrated 96% accuracy for even short 0.25 s samples to correctly classify emergency sirens.

A big challenge in human–vehicle interaction is the TOR from level-3 driving to manual control considering the driver’s situational awareness and distraction levels wherein voice assistants can be useful. Studies have been performed for the time taken by the driver to take control once a TOR has been issued with and without a VA.^[184] The VA offered a conversational discourse on traffic situation, infotainment suggestions, and calendar event reminders. The studies were conducted in a driving simulator and showed that VA helped in a timely takeover by 39%. Consequently, other research has also evaluated the use of conversational VA for takeover requests during level 3 automated driving.^[185] The study suggested that a simple countdown-based interface ranked highest in usability and perceived acceptance but offered least engagement and they proposed design guidelines for dialogue-based interaction for TOR.

VUI has inherent issues due to its sequential and temporal nature. Turn-taking issues are quite common wherein the users are not sure if the VUI is listening or preparing to respond.^[186,187] Users also need to rely on short-term memory which is inherently ephemeral.^[188,189] This reliance on short-term memory can be problematic during driving tasks when the driver needs to multitask between primary tasks and tertiary tasks. Furthermore, users complain of lack of control during voice interaction such as accessing particular menus and so

on.^[188–190] This motivates the use of multimodal interaction which is discussed in Section 4.3.

4.2.2. Display-Based Interaction

As discussed in Section 3.1, a central display unit is commonly available in modern vehicles wherein the inputs can be with traditional button and knob-based, touchscreen-based interaction, voice-interaction or gesture-based interaction.^[191] However, as they typically demand visual attention, HUDs and HMDs have been designed that fuse the informative virtual elements with the real scene and prevent the drivers to take eyes off the road. The HUDs are typically used during high-speed driving so that the driver need not take his eyes off the road while maneuvering through traffic. As one may imagine, for HUDs and HMDs to fix objects or cues to the real world scene, head pose tracking and registration is necessary. The registration method enables to fix virtual objects/text onto the real-world scene in a precise manner. More concretely, the following are necessary: 1) the absolute pose of the vehicle, 2) the head and eye relative position and orientation with respect to the vehicle coordinate system, and 3) identify and track objects (such as vehicles) outside the vehicle.^[192] A detailed description on head pose estimation is provided in Section 1.1.3. Registration algorithms need to be real time to prevent visual latency issues, adapt for vehicle vibrations, rapid head or eye movements and occlusions from the real scene. A real-time registration algorithm was proposed with an average registration time of 0.0781 s for AR HUD applications, thus preventing visual-latency issues.^[30] Vehicle vibration can cause issues for registration and a method was proposed for hiding virtual objects from HUD during large vibrations.^[193] A comprehensive review on registration methods for AR-based systems is provided in the work by Jiang et al.^[194] AR HUDs also present unique lighting and color blending challenges. As in automotive applications, there is a high variability of ambient light during daytime and nighttime driving requiring compensation in the AR HUD to be visible to the drivers. To tackle this, an active strategy was designed which samples the background scene and ambient light to adaptively adjust the brightness and color of the displayed AR images.^[195] Furthermore, in designing display elements, clutter plays a crucial role which can negatively affect usability if badly designed. Furthermore, accidental occlusions in HUDs can cause accidents wherein the virtual object is placed directly in the line of sight of the driver and occludes the real scene in front.^[192] Therefore, robust registration algorithms are necessary for deploying immersive and informative HUD/HMDs. As mentioned in Section 3.1, autostereoscopic displays can be used to provide display information in passive 3D format. Alerts and other necessary information can be provided through autostereoscopic displays as 3D information are known to capture human attention faster.^[196] For instance, studies have been performed to investigate the use of autostereoscopic displays in vehicle dashboards and found that users performed better in secondary tasks with 3D-like displays.^[196] The studies provided guidelines for incorporating S3D displays in vehicles. There is also a trade-off as complex displays are more attractive and immersive for drivers but increase distraction from primary task. Therefore,

proper design considerations are required while incorporating autostereoscopic displays.

4.2.3. Haptic-Based Interaction

Haptic perception involves haptic sensing as well as haptic feedback. The haptic modality is a combination of tactile and kinesthetic modes.^[197] For humans, tactile perception relates to understanding information from the skin, whereas kinesthetic perception is the perception of body positions and movements. Unlike visual and auditory modalities, the haptic modality requires contact with objects and the environment. In particular, humans perform active exploratory actions to perceive object properties, thereby perhaps changing the state of the world.^[198] Traditional haptic inputs include buttons, sliders, and rotary devices present in the dashboard or center console that allows the user to control vehicular functions without taking eyes off the road. Novel haptic interfaces discussed in Section 3.2 can be used for haptic sensing and feedback within the vehicle by covering the interior surfaces such as dashboard, hand rests, seats with tactile sensors, and vibrotactile feedback actuators. In this section, we explore methods to process haptic data and in particular tactile data as well as techniques from providing haptic feedback in various use-cases.

Possible use-cases for interaction with the haptic modality in particular vibrotactile feedback in the vehicle are detailed as follows^[199]:

- 1) spatial information from around the vehicles can be communicated directly via the skin thereby reducing visual cognition load;
- 2) warning signals are ideally suited for tactile modality about immediate dangers;
- 3) communication of private information silently to passengers without disturbing other passengers;
- 4) coded information of all types such as car status can be communicated to the driver when the driver requests or when specific conditions are met (such as in-vehicle climate control); and
- 5) general: providing information of settings of switches, preferences, and so on.

Gesturing is a natural and intuitive mode of communication developed in humans from birth and spans across cultures, ages, and tasks.^[200] Gesture-based interfaces can be broadly classified as 1) systems involving the use of wearable sensors for instance accelerometers, RFID tags, or data gloves^[201]; 2) touch sensitive interfaces^[202]; and 3) noncontact technologies such as depth cameras, thermal imaging cameras, ultrasonic tracking, and so on^[203,204] or a multimodal combination. Depending on the type of tactile sensors used, a human performing various action such as patting, scratching, or specific encoded gestures will generate high dimensional time-varying signals. To capture useful information, signal processing methods and data-driven machine learning algorithms are used. Raw tactile signals need to be converted to corresponding force and pressure values for subsequent processing.^[205] There have been extensive studies for tactile perception in the domain of robotics^[206] which can be transferred to the automotive domain. For instance, robust tactile descriptors have been proposed to extract tactile information regardless of number of sensors, sensing technologies, types,

and duration of exploratory movements or gestures.^[207–209] The proposed tactile descriptors are inspired by Hjorth parameters^[210] and represent the statistical properties of the tactile signals in the time domain, i.e., activity, mobility, and complexity. Furthermore, the novel tactile descriptors were used to classify various actions and gestures such as scratching, tickling, rubbing, poking, stroking, punching, patting, pushing, and slapping as well as a combination of the gestures using a SVM classifier.^[211–213] It was also shown that the framework is invariant to different subjects performing the actions and to contact locations as well as provided dynamic cell allocation wherein only those cells of the distributed tactile sensor were utilized for the classifier to reduce computational complexity and improving accuracy. The experiments were conducted on a humanoid robot, but they can be potentially transferred to the automotive domain for sensorized vehicle interiors. Tactile sensors can also be used to learn to recognize and distinguish objects in contact as demonstrated in previous studies.^[214,215] Recognizing gestures can be used to actuate parts of the car during autonomous mode (level 4/5) such as turning the chairs to face the rear and socialize with other passengers, control various vehicle infotainment functions, and so on. Touch-based and in-air gesture interactions with a novel active armrest designed with capacitive proximity sensors were studied in Braun et al.^[65] The framework consisted of three distinct steps. First, by using distributed capacitive proximity sensors the posture of the human arm is estimated by using a thresholding technique. Second, single touch or multitouch of fingers is detected using a weighted average and an additional thresholding of raw data. Upon preprocessing, the data are fed into a SVM classifier for recognizing gestures. An empirical study was performed to observe the effects on in-vehicle touch-based interfaces and in-air/midair gesture-based interfaces on driver distraction and user experience in Graichen et al.^[216] Subjective data, such as acceptance and workload, and objective data, including glance behavior, were collected from the participants of the study. Participants rated their perceived safety as higher while using gesture-based interaction and they performed significantly fewer glances at the visual display and the glances were shorter.

Upon sensing, haptic feedback is used by the intelligent vehicle interface to communicate and interact with the users. As explained in Section 4.1, vibrotactile feedback is typically used as a form of implicit communication to the drivers. Vibrotactile feedback is very effective to warn the driver of potentially dangerous situations such as collision during parking,^[76] lane departure,^[217] overspeeding,^[218] and so on. Apart from warning signals, haptic feedback can also provide informative signals. For instance, a study has been performed to provide navigation assistance using a wearable belt with eight tactors or devices providing vibrotactile feedback.^[219] The tactors triggered differently depending on the distance to a turn. The authors show that such a device can reduce visual cognitive load of following navigation instructions. Similarly, a 5x5 haptic feedback matrix was embedded on the seat and deployed for providing directions for navigation in Hwang et al.^[220]

Haptic interaction may not be the best modality in all use-cases as suggested in the study by Gaffary and Lécuyer.^[197] For instance, haptic feedback alone is prone to errors in navigation tasks as demonstrated in the study by Nukarinen et al.^[221]

Although alerts and controls could be provided with this modality, detailed information or instructions cannot be provided which is more effectively done using auditory or visual modality. Furthermore, careful design consideration when a driver should be provided with tactile feedback as improper use can be potentially startling and dangerous while driving. Therefore, there have been studies done wherein haptic modality is used in conjunction with other modes such as visual and auditory for effective communication.^[222–225] This is further explored in Section 4.3 on multimodal interaction.

4.3. Multimodal Interaction

Multimodal systems in user interaction may be defined as “those that process two or more combined user input modes—such as speech, pen, touch, manual gestures, gaze, and head and body movements—in a coordinated manner with multimedia system output”.^[226] As with monomodal interaction, multimodal interaction can have multiple inputs and outputs. Multimodal interaction can offer drivers different methods of interacting with the vehicle depending on the driving situation and the cognitive state of the driver. Furthermore, the drawbacks of any single modality can be compensated using another modality. One modality may even correct or verify the outputs of another modality. Various ways of incorporating multimodal inputs in automotive domain can be detailed as follows^[18]:

- 1) Temporally cascaded modalities: two or more modalities temporally sequenced such that the partial information which is recognized from the earlier modality is able to constrain the possible interpretations of the later modality. For instance, if a driver provides a speech input to change the menu on the head-down display screen and immediately uses the central control knob to manipulate it, the intelligent vehicle assistant (IVA) already can provide the intended menu as it had recognized the earlier command.

- 2) Redundant modalities: it is a special form of cascaded modality wherein given multiple modes of interaction; each modality is available at each step. For instance, navigation systems that can be controlled by touchscreen inputs and equivalently by voice inputs.

- 3) Fused modalities: multiple modalities can be fused as part of a single interaction step. For instance, pointing at a structure while driving and asking “what is that” to the virtual assistant fuses gesture, speech, and gaze.

Speech modality is often used in conjunction with visual or tactile feedback as a multimodal interaction method.^[227] For example, a system of interacting with the voice user-interfaces with gestures for fine-grained manipulation and easy-to-undo actions has been designed in the work by Pfleging et al.^[228] The study compares the voice+gesture multimodal input with physical buttons as baseline and results show that although multimodal input is slower, it results in similar driving performance and reduces the visual demand. Similarly, multimodal output of speech and visual aids for effective user interaction was investigated in the work by Braun et al.^[229] The voice user-interface (VUI) is augmented with visual texts and icons for increasing the effectiveness of interaction and minimizing distractions. Their studies conducted with 64 participants in a driver

simulation setup resulted in the following conclusions: 1) text summaries help drivers remember facts and improve the user interface, but they can also cause distraction; 2) keywords reduce cognitive load and have a positive impact on driving efficiency, and 3) the use of icons enhances the appeal of the user interface. Likewise, voice commands were studied in conjunction with haptic buttons and knobs on the steering wheel to navigate and correct the dictated sentences.^[229] The study found that using visual feedback, voice commands, and manual interactions together may cause significant distraction. Considering haptics and voice-based interaction, a study was performed with a voice+tactile framework wherein VUI is augmented with high-resolution tactile outputs.^[230] The study had four voice+tactile schemes: Status Feedback, Input Adjustment, Output Control, and Finger Feedforward. User studies showed that the proposed framework improved the VUI efficiency and its user experiences without incurring significant additional distraction overhead on driving. Therefore, the study concluded that multimodality can help reduce driver distraction.

Considering other multimodal inputs, gestures have been used in conjunction with visual modalities such as head-up displays to control menus.^[231] Another study compared interacting with IVIS with gestures on steering wheel, touch on central dashboard, and speech.^[232] The study found no statistically significant improvement in performance with any modality while the touch modality required the least time to completion. Considering techniques for processing multimodal data, DNNs are frequently used for multimodal fusion such as to combine the gaze, head pose, and finger pointing gestures for object selection.^[233] Similarly, another framework demonstrated use-cases for integrated multimodal modulation of in-vehicle functions using either a single modality (speech, look, or gesture) or a combination of two or more.^[234] Having multiple modalities can help the driver to choose which modality to use for communication that is least affected by environmental influences.^[235] Multimodal inputs can be used for controlling vehicle functions in addition to selecting a particular task or object. For instance, a mixture of three modalities, voice, gaze, and movements, was used to pick vehicle objects, such as side mirrors or windows, and then control these objects with gestures or speech in the study by Neßelrath et al.^[236]

In-air gestures, i.e., those which do not require contact with a surface and controlled through multimodal interfaces, are increasingly available in commercial vehicles such as the BMW iNext.^[237] The goal of gesture-based interaction (GBI) is to reduce the cognitive workload, reduce the demand for visual attention thereby increasing safety.^[238] As with designing interactive interfaces, it is necessary to ensure the GBI mechanism aid in reducing driver distraction. As an example of midair gestures, an RGB-D sensor was used to track hand gestures in Riener et al.^[44] The participants controlled a display screen through in-air gestures by pointing actions. Their studies showed that the system is capable of discerning static finger and hand gestures than dynamic gestures. The gestures are performed while placing the hand on the gear-shift area, thereby reducing the safety concern. In a similar way, a study was performed to evaluate the degradation of driver performance using in-air gestures versus touch gestures to control an infotainment system (secondary task) while performing the driving task (primary task).^[239]

They performed the The Lane Change Test (ISO 26022 standard^[240]) which aims at measuring the degradation of human performance with respect to a certain primary task while conducting a secondary task. The result served as an estimate for the demand of the secondary task. The study consisting of 17 participants found comparable results for driver performance degradable for both types of interactions. However, as all participants were experienced with touch-based interactions and had no experience with in-air gestures, the study showed promising avenues for in-air gesture control. Eye-gaze and midair pointing gestures were demonstrated to be the preferred mode of recognizing points of interests for participants in a study.^[241] Furthermore, the participants regulated the vehicle speed (primary task) while interacting using pointing gestures, thus compensating for distraction.

There is a need for large-scale multimodal datasets to foster research focus toward development of fusion algorithms and for comparing various proposed methods. Multimodal interaction has shown great potential and future directions will involve developing robust algorithms for detecting driver distraction, emotion, and activity using multitude of sensors increasingly present in the cars.

5. Discussions: Challenges and Outlook

HVI research is still in its infancy in comparison with similar areas such as wearable systems and home interiors. In this regard, the overview of various sensing modalities and techniques for in-vehicle interaction provided here is just the beginning of this rapidly emerging field. Like any newly emerging field there are multiple challenges intertwined with huge opportunities for future research and some of these are described later. While some of the challenges (e.g., communication, energy, etc.) are common with areas such as wearable systems, others (e.g., interior designs) are unique to automotive sector.

5.1. Empirical Studies for HVI

Many of the technological solutions developed so far are used for various HVI studies (e.g., driver distraction and emotion recognition) that are typically conducted in driving simulator environments,^[242] open source annotated datasets, or controlled outdoor environments with preplanned routes. Although in controlled environments, it is possible to elicit emotions by providing relevant triggers, such studies are not always representative of the real emotional or cognitive state in real-world conditions.^[5] Furthermore, many of these studies are not diverse enough to cover different cultural and geographical situations. Therefore, “universal” annotated datasets which are unbiased to geographical, environmental, cultural, regional, and traffic signal variations are necessary to test various algorithms and technologies developed for HVI. In connection to this, the issue of annotation errors in tasks such as emotion recognition from facial expressions which is subjective and is dependent on the annotator’s mental state, cultural and geographical influences are relevant.^[243] Datasets having different annotation biases cannot be effectively merged during the training process and models evaluated on cross-datasets demonstrate deteriorating

performance.^[244] Some works have designed frameworks that can robustly learn from various datasets with differing annotator preferences.^[245] Methods such as Expectation Maximization (EM) have been used to identify the latent ground-truths from different annotators.^[246,247] Another issue as noted in Li and Deng^[243] is the data imbalance due to practical data collection reasons: annotating and collection of smiling emotion is easier than eliciting emotions such as disgust or anger. A possible solution may be balancing the class distribution during preprocessing using data augmentation. Research toward developing models robust to annotation bias and labeling errors deserves attention. Moreover, majority of the recognition studies are performed with initial data collection and postprocessing is done to compare different methodologies. In this regard, there is need to study the online recognition systems with real-world experiments. To this end, robust and certifiable perception algorithms and decision-making frameworks are needed so that failure can be easily noted.^[248]

5.2. Robust Multimodal Perception and Interaction

With a multitude of sensors present within the car, the research on effectively combining different complementary sensing information becomes necessary. As an example, with novel haptic interfaces being integrated into the vehicles, the visuo-haptic fusion techniques need to be designed to utilize complementary qualities of vision and haptic data.^[206,249] In this regard, the methods developed in the robotic manipulation domain can serve as reference. Furthermore, cross-modal transfer between vision and touch sensory modalities can bring redundancy in the system, making it more reliable, especially when one sensing modality is unavailable.^[250–254] For instance, pose estimation within the vehicle can be performed using multimodal fusion approach of visuo-haptic data. It is important to formalize a representation for visual and haptic data considering the dense and sparse nature of these modalities. In this regard, a novel active approach based on Bayesian filtering has been designed which can handle sparse tactile data in addition to dense visual data for robustly estimating object pose.^[255] Multimodal sensing further provides the opportunity to detect faulty or erroneous data from one sensing modality using another modality. Such corrupted information may be rejected for any inference based on sensor inputs. In addition, alternative sensing modalities can also correct the faulty modality using the built-in redundancy of the data. The active selection of the correct sensing modality for a

given task is necessary as depending on the nature of the task; a certain modality may perform better in monomodal fashion rather than through multimodal fusion. Therefore, appropriate sensor selection is critical for a safe and reliable vehicular system.^[14] Another novel research path is interior and exterior sensor fusion. This can assist in predictive support wherein internal sensors monitor driver state, distractions, and fatigue levels and fuse the data from exterior sensors which can detect obstacles and pedestrian intention and proactively inform the driver preventing a late intervention.^[256] Perception errors due to sensing inaccuracies or algorithmic errors can be corrected on-the-fly through human–vehicle teaching as well. Through HVI, the vehicle can learn from humans and correct its internal models.^[257]

5.3. Functional Vehicle Interior Designs

Starting from level 3, the role of the driver can change gradually to a mere passenger and responsibilities shift from focusing on the traffic and vehicle control to more hedonic and entertainment oriented. For instance, in level 4 the steering wheels and accelerator and braking pedals will be minimally used when the driver needs to take over control and in level 5, they need not be present at all. Therefore, new possibilities of designing the interiors of the car will emerge. The seats no longer need to face the front; they could even be turned to face each other as shown in **Figure 7**. Such configuration (with rotating seats in front-facing and rear-facing) has been investigated for TOR task in a driving simulator.^[258] Numerous entertainment units can fill up the commute time through large screen displays, holograms, or workstations. Novel designs can also foster socializing with other passengers in a natural way as well as reconfigurable interiors wherein the interiors can be modified into various configurations to satisfy the personal preferences of the passengers. New interior interfaces also need to clearly communicate to the users the intentions of the AV as well as guide the users who will be unfamiliar with the interfaces to take over control and so on. Therefore, the interior designs will be crucial to foster trust and acceptance of AVs by the people. Moreover, novel interior designs will also foster other challenges such as haptic sensing surfaces that can adapt to reconfiguration, wiring considerations (or wireless communication), actuators capable of performing smooth reconfiguration by understanding the pose and intent of the users, and so on. Furthermore, as an increasing number of sensors are integrated into the vehicle, interesting questions on placement of sensors, cost, and redundancy of sensors will

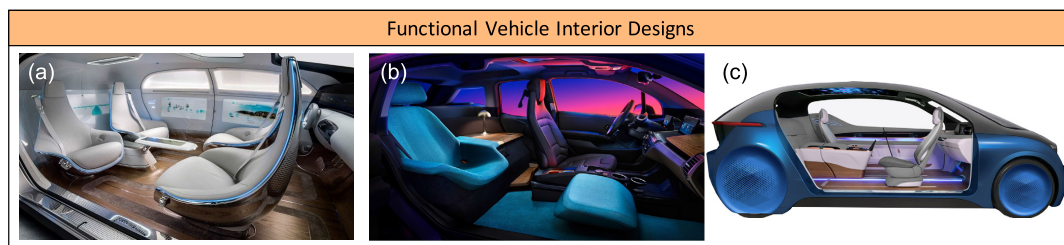


Figure 7. a–c) Examples of functional interiors for concept vehicles. Part (a): Reproduced with permission.^[306] Copyright 2021, Daimler AG. Part (b): Reproduced with permission.^[307] Copyright 2021, BMW Group. Part (c): Reproduced with permission.^[79] Copyright 2021, Yanfeng Automotive Interior Systems.

emerge. As various sensor setups are designed in a very early stage of vehicle development, it is challenging to test the different types of sensors, sensor positioning or number of sensors in real-world environments. For sensor positioning, number of sensors or even make the design future-proof. Therefore, simulations can play a crucial role in terms of designing sensors and their positioning. In this regard, the work involving simulation-based evaluation of arbitrary sensor setups for environment perception is worth noting.^[259] Similar research efforts toward simulation of interior perception are needed.

5.4. Communication Protocols

Traditional communication systems such as local interconnect network (LIN),^[260] controller area network (CAN),^[261] FlexRay,^[262] and CAN with flexible data rate (CAN-FD)^[263] have been used for in-vehicle communication over the years. The modern vehicle electrical/electronic (E/E) architecture is focused on automotive Ethernet (1000Base-T1 and 100Base-T1).^[264] Ethernet is a key enabler for level 2/ level 3 autonomous driving functionalities as well as for intelligent vehicle assistants. It appears that the automotive Ethernet is the future of in-vehicle networking.^[265] This is because 1) It directly supports service-oriented architectures (SOA); 2) AVs use high compute systems as well as diverse sensing devices which work on IP-based networking. Therefore, high-bandwidth, full-duplex Ethernet is required to exchange data between them; 3) in traditional CAN-based networking, the essential information of the CAN bus needs to be available to other services outside the car as well where the receiver would face difficulty understanding the vendor-specific CAN bus data. Ethernet protocol and IP addressing are well established also outside the vehicles. Comparison between the various intravehicle protocols is shown in Table 3.

New technologies such as 5 G/ 6 G offering ultralow latencies and ultrahigh reliability can be used for V2V/V2X networks.^[266] On the lines of ultralow latencies and ultrahigh reliability, the notion of Tactile Internet has emerged recently. The IEEE P1918.1 working group^[267] defines the Tactile Internet as “a network, or a network of networks, for remotely accessing, perceiving, manipulating, or controlling real and virtual objects or

processes in perceived real time.” Particular applications include platooning or cooperated automated driving wherein vehicles in a platoon autonomously follow each other at a close distance, all driven by a common shared leader. The leading vehicle may be manually controlled by a professional driver, autonomously driven or teleoperated.^[268] Thus, platooning is a safety critical application wherein follower vehicles rely on local perception information as well as V2V information transmitted wirelessly from other vehicles as well as the leader vehicle. The state-of-art networking protocols are still lacking in terms of low latencies and reliability guarantees that the tactile Internet is aiming to tackle.^[269] Detailed reviews on tactile internet can be found elsewhere.^[269,270] In-vehicle networks are also susceptible to malicious adversarial attacks detailed in Section 5.5.

5.5. Security, Privacy, and Trust

Protecting in-vehicle sensors from adversarial cyberattacks is crucial for the safety and privacy of the vehicle users. As the sensors and most vehicular functions such as steering, sensors, and so on are controlled by electronic control units (ECUs), they are susceptible to adversarial attacks. For instance, GPS signals can be attacked by spoofing or jamming the signals. GPS signals can be jammed by sending signals at the same frequency.^[271] Once jammed, the attacked may spoof the GPS signal by locking onto the GPS detector with a spurious signal which can then provide abrupt changes to the vehicle's ego position.^[272,273] Some studies to investigate possible attacks on sensors such as LIDAR, RADAR, and cameras have brought to fore some interesting facts. For example, fake obstacles can be presented to a LIDAR by using a transceiver to receive the laser pulse sent by the LIDAR and delay the response pulse back to the LIDAR.^[274,275] Similarly, interior sensors such as microphones and cameras are also susceptible to potential attacks to the privacy of the passengers.^[276]

In-vehicle communication channels as well as V2V and V2X communication streams can also be susceptible to adversarial attacks. For example, remote attacks could be performed without any alteration to the vehicle and control of critical functions such as braking and steering. A similar case led Fiat-Chrysler

Table 3. Comparison between selected intravehicle networks.^[308,309] Note: the power consumption is based on average values for computed for a typical transceiver device from each protocol.^[310–313]

Protocol	Bandwidth [bits s ⁻¹]	Max efficiency	Typical topology	Power requirement [mW]	Relative cost	Selected application
LIN	<20 K	51.6%	Linear bus	45–75	Low	Battery monitoring, window control, steering wheel, temperature sensor, sunroof control, alternator module
CAN	100–500 K	59.6%	Mostly linear bus	≈50	Low/ Medium	Engine controller, transmission unit, seat module, climate control, electrical stability control, headlamp assembly
FlexRay	5–10 M	97%	Linear bus, star or hybrid	≈50	High	Steering angle sensor, safety RADAR, all-wheel drive, dynamic suspension control, throttle control, network backbones
Ethernet	100 M	97.5%	Point-to-point	≈150 + 50 ^{a)}	Medium	Cameras, LIDAR, RADAR, entertainment unit, navigation system, audio module, amplifier control, wireless electronics connector, network backbones

^{a)}For Ethernet, each port adds up to 50 mW of power consumption.

Automobiles (FCA) to recall 1.4 million vehicles.^[277] Similarly, the CAN protocol is vulnerable to malicious attacks through injection of spurious packets into the bus.^[278] For instance, a type of denial-of-service (DoS) attack, called bus-off attack, may exploit the error handling scheme of the CAN bus to shut down or disengage ECUs.^[279] On the contrary, CAN-FD provides a larger payload that could be used for encryption of the data. Blockchain frameworks can also be used for protecting in-vehicle networks.^[280] An open challenge for blockchain-based vehicle security is the consensus mechanism, i.e., the election of peers for the consensus algorithm^[280] as it directly impacts the integrity of the data in the network. Another challenge for blockchain is to integrate with existing in-vehicle protocols such as CAN and LIN.

As detailed in Section 4, DNNs are commonly used for perception from raw sensor signals for interior as well as exterior of the vehicle. DNNs are very vulnerable to adversarial examples as demonstrated very recently by OpenAI^[281] wherein, among other examples, a state-of-art DNN classified an apple incorrectly just by placing a sticker with “iPod” written on it. These attacks are called typographic attacks. Similarly, an AV was caused to drive recklessly by deliberately generating toxic signs on the road by an attacker.^[282] Defense against adversarial attacks is a rapidly progressing research field and explainable or interpretable learning models may prove to be a way forward for robust perception and decision-making.

In-vehicle systems are also susceptible to breaches of privacy. For example, exterior cameras may capture sensitive videos and data of pedestrians as well as confidential areas such as militarized zones.^[283] Similarly, interior cameras and microphones can always listen in and observe the driver and passenger activities. Even if these sensors are turned off, vehicle’s trajectories from GPS data can reveal information such as work and home addresses.^[284] Haptic interfaces monitoring vital signals can reveal sensitive health-related information to any potential attack. More research is required for data privacy and protection for AVs.^[285]

Overreliance and trust on immersive interactive technologies may prove to be detrimental during assisted driving. Studies have shown that humans are incapable of effectively focusing on multiple information channels simultaneously.^[286] For instance, drivers may become overreliant on AR cues provided through HUDs or HMDs such as taking the next turn and prioritize it over real-world cues such as sudden vehicle stoppage in front of the vehicle.^[287] Research toward trust on the virtual assistants in the car requires attention as over trust can result in missing real-world cues and mistrust may result in the driver ignoring the interface completely, thus resulting in no interaction.^[192] Research on transparency and explainability of AVs is also crucial to improve the trust and adoption of AVs in the future.

5.6. Data Processing and Energy Challenges

5.6.1. Data Processing

With a multitude of sensors being integrated into the vehicle, the amount of information collected and processed by the vehicle increases rapidly. In a study conducted by Intel in 2016 showed that AVs will generate up to 4000 GB of data every day and each AV will generate the same amount of data as 3000 people.^[288] For

instance, the amount of data generated by cameras (40 MB/s), LIDAR (70 MB/s), RADAR (100 KB/s), GPS (50 KB/s), and Ultrasonic (100 KB/s), whereas novel interior sensors are yet to be benchmarked for the amount of data collected. Powerful compute hardware is included in vehicles to handle the large amount of sensor data produced and perform inference and control actions. For instance, graphic processor unit (GPU) such as NVidia DRIVE PX2 and NVidia DRIVE AGX are capable of performing 30 and 320 trillion operations per second (TOPS) while utilizing 60 and 300 W power, respectively.^[285] Furthermore, high-performance compute hardware raises the cost of the overall system significantly. As a consequence, affordable edge computing devices that are low powered have seen research interest^[289] but require further attention. Moreover, new methods for handling the information flood are required and are an active area of research. Active perception can mitigate collection and processing of redundant data.^[290] Compute resources can be centralized wherein the compute hardware and associated power source lie in the same space within the vehicle but does not provide fail-safe redundancies for the sensors that are distributed around the vehicle. In contrast, by distributing compute infrastructure in the vehicle can increase robustness to failures and provide redundancy for safety critical features such as object detection and steering during level 3–5. Additional communication channels for distributed computing may lead to increased costs and weight of the vehicle due to increased wiring which can be potentially tackled through wireless communication.^[291] Distributed computing framework for AVs relies on middleware software stacks such as Data Distribution Service (DDS),^[292] Robot Operating System (ROS),^[293] Automotive Open System Architecture (AutoSAR),^[294] and so on for seamless integration of different functionalities, services, and compute hardware. However, automotive OEMs may prefer to use proprietary middleware as part of their autonomous driving software stack which can decrease software reusability and vendor specific software modules.^[285] Therefore, flexible software architecture capable of using vendor-specific components, open source software as well as different real-time processing, is required.^[295] A robust, open-source real-time vehicle operating system remains to be an open challenge. Furthermore, for real-time functioning, various data sources from multimodal sensors require proper time synchronization which can be potentially problematic due to the different relative frequencies of the sensors. Time synchronization using network time protocol (NTP) can lead to timestamp differences of up to 100 ms.^[285] A comprehensive review on data processing and computing challenges for autonomous driving can be found elsewhere.^[285]

5.6.2. Energy Challenges

Typical battery capacity of full battery electric vehicles (BEVs) ranges from 80 to 100 kWh. For fully AVs, the auxiliary electrical load requirement is expected to increase by few kilowatts to power the various on-board electronics, sensors, and computation units such as GPUs. Thus, the range of full autonomous electrical vehicle would be reduced compared with the standard BEV. For instance, the heating, ventilation, and air condition system (HVAC) is one of the largest auxiliary loads on the BEV.^[291] Sensors and computing units are other auxiliary electrical loads

which are connected to the battery power source via a DC/DC converter connected to a 12 V bus. For a minimal number of sensors and compute for an AV such as LIDAR, RADAR, IMU, cameras, and NVidia GPU, the power requirement is estimated to be approximately 200 W.^[291] This does not consider the interior sensing stack, display units, and infotainment systems. A study showed that energy consumption of connected and AVs is expected to increase the current energy consumption up to 15%.^[296] Therefore, increase in autonomy without accounting for energy efficiency will reduce range and increase costs for the consumers, thereby reducing the potential adoption by consumers.

Considering renewable sources of energy, recent studies have shown that novel low-cost, thin film and flexible solar cells integrated into all upward facing parts of the vehicle may provide a viable solution for battery range extension.^[297] Simulation studies with such solar cells showed up to 17.5% reduction in annual net vehicle energy usage and average range extension of 47 km/day in Detroit on an average sunny day.^[297] Similarly, a novel electronic tactile skin designed with miniaturized solar cells showed the potential to generate more than 100 W if covered over an area of 1.5 m².^[298] In addition, vehicle-to-grid (V2G) technology will enable electric vehicles to provide energy back to the grid in case of increased demands as well as to store and discharge electricity produced from renewable energy such as solar or wind which can fluctuate during the day.^[299]

Energy can be stored in vehicles in a centralized or distributed form. Apart from redundancy in communication wiring as seen in Section 5.6.1, redundancy for wiring for energy distribution is also critical. A typical nonautomated sedan will have wiring up to 3640 m and additional communication and power wiring due to sensing such as LIDAR and RADAR may be up to 120 m.^[291] A central power source would require lesser and simpler wiring architecture but compromises on fail-safe redundancy of power failures to any sensing, compute, or motor units. On the contrary, distributed power storage with small storage units strategically placed around the vehicle can offer redundancy to power failures in wiring to any device. However, this would increase the length of wires and cabling and also the overall weight of the vehicle. Therefore, a trade-off needs to be struck for redundancy and power efficiency. Research in this field is nascent and novel techniques for distributed energy harvesting and storage are required to account for increase in auxiliary electrical loads.^[94] With the growing number of BEVs, environmental sustainability issues have been raised with increasing use of batteries. Sustainability must be a criterion at every stage of vehicle production and in particular, biocompatible and biodegradable materials could be used for energy harvesting and storage.^[94]

Therefore, there is a clear challenge for energy and data efficient sensing and compute devices within the car. To have natural interaction and increased autonomy, edge compute devices require high computational capabilities. Intel, NVidia, and Qualcomm are building processing units optimized for self-driving cars with high compute and low power consumption.^[300] Traditional GPUs that are used to execute DNNs are exceptionally power and data hungry which can drain out the vehicle's battery whereas relying on cloud computing can increase lag and latency issues in safety-critical processes such as driving. Recent studies have shown that neuromorphic technologies enable energy efficient edge computing possible by using a

thousand times less power than a traditional GPU.^[301,302] For instance, spiking neural networks (SNNs) running on an embedded neuromorphic hardware have been used to recognize voice commands 0.2 s faster than a GPU alternative while consuming much less power.^[301] Low-powered edge computing devices implemented on NVidia Jetson TX1 module have been shown to support most AV functionalities while consuming as less as 11 W of power in comparison to high powered GPUs that consume around 300 W.^[289]

Similar to neuromorphic computing devices, low-powered and neural-like sensing devices have also emerged to tackle the data processing and energy requirement problems. Although it is promising to consider the entire vehicle interior covered with large-area flexible tactile sensing e-skin, the information processing of a large amount of data emerging from the distributed tactile sensing has remained a challenge.^[303] Similarly, powering such large area tactile sensors can also be a constraint for future battery electric vehicles. There has been research for distributed local processing of tactile data performed in situ to release the computational load from the central processing compute units. For instance, a novel hardware-implementable neural network based on neural nanowire field effect transistors (ν -NWFETs) has been proposed for distributed local tactile neuromorphic processing mimicking a human skin.^[304] Essentially, the possibility of printing nanowires^[305] shows potential to fabricate large area, conformable tactile sensors capable of local distributed computations to be covered on curved surfaces.^[39]

Research toward novel energy management systems, high compute hardware with low energy consumption, and low-powered or self-powered sensing devices are necessary for energy-efficient AVs.

6. Conclusions

With increasing levels of automation and intelligence in the vehicle, the scope for interaction between the human and the vehicle is also growing. While there has been tremendous research in the field of environment perception and control for autonomous driving, the growing need for in-vehicle interaction is also leading to more sensors being integrated within the car. Legacy display interfaces and haptic devices such as buttons and knobs are being replaced by novel multi-modal sensing devices. In this article, we have attempted to provide a review of the state-of-art of human-vehicle interaction in terms of automotive sensing technologies and associated methods and techniques for providing natural and intuitive interaction. The role and need for such interactive interfaces are outlined. Recent technologies which enable interaction are reviewed. State-of-the-art methods which enable a vehicle to understand the implicit contextual cues of communication as well as explicit communication modes such as speech and gestures are discussed. Furthermore, the advantages and disadvantages of using any particular modality for interaction have been detailed as well as methods for multimodal fusion have been outlined. Finally, we discussed the current challenges in human-vehicle interaction research and raised several questions such as: How can AVs understand contextual cues using multiple sources of sensing data? How can we develop intelligent vehicle assistants that are not only effective to interact but are also empathetic

toward the users? How can AVs analyze passenger behavior and activities while respecting privacy and being nonintrusive? How can novel interior be codesigned with sensing and interaction modes in mind? How can AVs change the internal structure or functions to minimize occupant stress and increase productivity or comfort? How can human-machine interfaces (HMIs) in AVs foster trust through transparency and explainability of actions and intentions? What kind of communication protocols are necessary to keep up with growing number of sensors and information generation in the vehicle? How can AVs achieve energy efficiency whilst consuming battery power for motors, sensors, intelligent interiors, and so on? We believe that future interactive AVs should focus on tackling these questions.

7. Appendix

Table of acronyms used in this article.

Acronym	Explanation
AV	Autonomous vehicle
HVI	Human-vehicle interaction
HRI	Human-robot interaction
FSD	Full self-driving
OEM	Original equipment manufacturer
SAE	Society of Automotive Engineers
V2V	Vehicle-to-vehicle
V2P	Vehicle-to-pedestrian
V2X	Vehicle-to-Everything
V2G	Vehicle-to-grid
V2I	Vehicle-to-infrastructure
IVIS	In-vehicle infotainment system
IVA	Intelligent vehicle assistant
TOR	Take over request
HDD	Head-down display
HUD	Head-up display
HMD	Head-mounted display
AR	Augmented reality
FTIR	Frustrated total internal reflection
HAD	Highly automated driving
S3D	Autostereoscopic display
LCD	Liquid crystal display
RGB-D	Red green blue-depth
LIDAR	Light detection and ranging
RADAR	Radio detection and ranging
IR	Infrared
ToF	Time-of-flight
PTZ	Pan-tilt-zoom
IR-UWB	Impulse radio ultrawide band
LED	Light-emitting diode
EEG	Electroencephalography
ECG	Electrocardiography
EOG	Electrooculography

Acronym	Explanation
EDA	Electrodermal activity
EMG	Electromyography
GUI	Graphical user interface
IMU	Inertial measurement unit
VR	Virtual reality
MR	Mixed reality
HMI	Human-machine interface
NHTSA	National Highway Traffic Safety Administration
GPS	Global positioning system
CNN	Convolutional neural network
DNN	Deep neural network
SVM	Support vector machine
PERCLOS	Percentage of eye closure
kNN	k-nearest neighbors
PCA	Principal component analysis
LDA	Linear discriminant analysis
BNN	Bayesian neural network
GMM	Gaussian mixture models
MNR	Multinomial regression
VUI	Voice user interface
VA	Voice assistant
RFID	Radio-frequency identification
GBI	Gesture-based interaction
LIN	Local interconnect network
CAN	Controller area network
CAN-FD	Controller area network with fixed data rate
E/E	Electrical/electronic
SOA	Service-oriented architecture
ECU	Electronic control unit
DoS	Denial of service
TOPS	Trillion operations per second
DDS	Data Distribution Service
ROS	Robot Operating System
AutoSAR	Automotive Open System Architecture
NTP	Network time protocol
BEV	Battery electric vehicle
GPU	Graphics processing unit
HVAC	Heating, ventilation, and air condition system
SNN	Spiking neural network
NWFET	Neural nanowire field effect transistor

Acknowledgements

The authors would like to thank Ms. Alicja Slepowska and Mr. Anirvan Dutta for assistance with the figures in the article. Furthermore, the authors would like to sincerely thank Mr. Michael Gentner, Mr. Xavier

Anzuela Recasens, and Dr. Hans-Jörg Vögel for the fruitful discussions. This work was supported, in part, by the European Commission via iNavigate (grant no. 873178), SmartNets (grant no. 860949), and Intuitive (grant no. 861166).

Conflict of Interest

The authors declare no conflict of interest.

Keywords

autonomous vehicles, human–vehicle interaction, interactive interiors, multimodal interaction

Received: June 16, 2021

Revised: August 9, 2021

Published online: October 3, 2021

- [1] Sae j3016 201806: Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles, https://www.sae.org/standards/content/j3016_201806/ (accessed: January 2021).
- [2] M. A. Goodrich, A. C. Schultz, *Human–Robot Interaction: A Survey*, Now Publishers Inc, Hanover, MA, USA, **2008**.
- [3] F. Biondi, I. Alvarez, K.-A. Jeong, *Int. J. Human–Comput. Interaction* **2019**, 35, 932.
- [4] Y. Dong, Z. Hu, K. Uchimura, N. Murayama, *IEEE Trans. Intell. Transp. Syst.* **2010**, 12, 596.
- [5] S. Zepf, J. Hernandez, A. Schmitt, W. Minker, R. W. Picard, *ACM Comput. Surveys (CSUR)* **2020**, 53, 1.
- [6] A. Sahayadhas, K. Sundaraj, M. Murugappan, *Sensors* **2012**, 12, 16937.
- [7] G. Borghini, L. Astolfi, G. Vecchiato, D. Mattia, F. Babiloni, *Neurosci. Biobehav. Rev.* **2014**, 44, 58.
- [8] X. Sun, S. Cao, P. Tang, *Appl. Ergonom.* **2021**, 90, 103238.
- [9] H. Detjen, S. Faltaous, B. Pfleging, S. Geisler, S. Schneegass, *Int. J. Human–Comput. Interaction* **2020**, 37, 1.
- [10] A. L. Kun, *Found. Trends Human–Comput. Interaction* **2018**, 11, 201.
- [11] A. Rasouli, J. K. Tsotsos, *IEEE Trans. Intell. Transp. Syst.* **2019**, 21, 900.
- [12] A. Rouchitsas, H. Alm, *Front. Psychol.* **2019**, 10, 2757.
- [13] C. Bila, F. Sivrikaya, M. A. Khan, S. Albayrak, *IEEE Trans. Intell. Transp. Syst.* **2016**, 18, 1046.
- [14] A. Sarker, H. Shen, M. Rahman, M. Chowdhury, K. Dey, F. Li, Y. Wang, H. S. Narman, *IEEE Trans. Intell. Transp. Syst.* **2019**, 21, 7.
- [15] L. Hobert, A. Festag, I. Llatser, L. Altomare, F. Visintainer, A. Kovacs, *IEEE Commun. Mag.* **2015**, 53, 64.
- [16] H. Onishi, in *2018 IEEE 88th Vehicular Technology Conf. (VTC-Fall)*, IEEE, Piscataway, NJ **2018**, pp. 1–6.
- [17] G. Geiser, *ATZ* **1985**, 87, 56.
- [18] C. Müller, G. Weinberg, *IEEE MultiMedia* **2011**, 18, 98.
- [19] Y. Du, J. Qin, S. Zhang, S. Cao, J. Dou, in *Int. Conf. on Human–Comput. Interaction*, Springer, Springer Nature, Cham Switzerland, **2018**, pp. 117–132.
- [20] H. H. van Huysduynen, J. Terken, B. Eggen, in *Proc. of the 10th Int. Conf. on Automotive User Interfaces and Interactive Vehicular Applications*, **2018**, pp. 247–257.
- [21] C. P. Janssen, A. L. Kun, S. Brewster, L. N. Boyle, D. P. Brumby, L. L. Chuang, in *Proc. of the 11th Int. Conf. on Automotive User Interfaces and Interactive Vehicular Applications: Adjunct Proc.*, **2019**, pp. 465–467.
- [22] R. Schroeter, F. Steinberger, in *Proc. of the 28th Australian Conf. on Computer–Human Interaction*, **2016**, pp. 25–29.
- [23] S. Diewald, A. Möller, L. Roalter, T. Stockinger, M. Kranz, in *Proc. of the 5th Int. Conf. on Automotive User Interfaces and Interactive Vehicular Applications*, **2013**, pp. 262–265.
- [24] Land rover: The world's first self-learning car, <https://www.landrover.com/experiences/news/self-driving-car.html> (accessed: January 2021).
- [25] H.-J. Vögel, C. Süß, T. Hubregtsen, E. André, B. Schuller, J. Härrig, J. Conradt, A. Adi, A. Zadorojniy, J. Terken, J. Beskow, A. Morrison, K. Eng, F. Eyben, S. Al Moubayed, S. Müller, N. Cummins, V. Ghaderi, R. Chadowitz, R. Troncy, B. Huet, M. Önen, A. Ksentini, in *2018 IEEE/ACM 1st Int. Workshop on Software Engineering for AI in Autonomous Systems (SEFAIAS)*, IEEE, Piscataway, NJ **2018**, pp. 11–15.
- [26] F. Lauber, A. Follmann, A. Butz, in *Proc. of the 2014 Conf. on Designing Interactive Systems*, **2014**, pp. 171–180.
- [27] Mercedes-benz vision eqs, <https://www.mercedes-benz.com/en/eq/concept-cars/vision-eqs/> (accessed: January 2021).
- [28] V. Charissis, S. Papanastasiou, *Cognit. Technol. Work* **2010**, 12, 41.
- [29] R. B. Wood, M. J. Hayford, in *Holographic Optics: Design and Applications*, Vol. 883, International Society for Optics and Photonics, Bellingham, WA **1988**, pp. 36–52.
- [30] Z. An, X. Xu, J. Yang, Y. Liu, Y. Yan, *IEEE Access* **2018**, 6, 43749.
- [31] K. Akaho, T. Nakagawa, Y. Yamaguchi, K. Kawai, H. Kato, S. Nishida, *Electr. Eng. Jpn.* **2012**, 180, 43.
- [32] R. F. Berriel, E. de Aguiar, A. F. De Souza, T. Oliveira-Santos, *Image Vision Comput.* **2017**, 68, 64.
- [33] W. Narzt, G. Pomberger, A. Ferscha, D. Kolb, R. Müller, J. Wiegardt, H. Hörtnner, C. Lindinger, *Universal Access Inf. Soc.* **2006**, 4, 177.
- [34] A. Christou, Y. Gao, W. T. Navaraj, H. Nassar, R. Dahiya, *Adv. Intell. Syst.* **2020**, 2000126.
- [35] A. Christou, R. Chirila, R. Dahiya, *Adv. Intell. Syst.* **2021**, in press.
- [36] L. Manjakkal, L. Yin, A. Nathan, J. Wang, R. Dahiya, *Adv. Mater.* **2021**.
- [37] E. S. Hosseini, S. Dervin, P. Ganguly, R. Dahiya, *ACS Appl. Bio Mater.* **2020**, 4, 163.
- [38] E. O. Polat, O. Balci, N. Kakenov, H. B. Uzlu, C. Kocabas, R. Dahiya, *Sci. Rep.* **2015**, 5, 1.
- [39] R. Dahiya, W. T. Navaraj, S. Khan, E. O. Polat, *Inform. Display* **2015**, 31, 6.
- [40] C. G. Núñez, W. T. Navaraj, E. O. Polat, R. Dahiya, *Adv. Funct. Mater.* **2017**, 27, 1606287.
- [41] C. W. Hsu, B. Zhen, W. Qiu, O. Shapira, B. G. DeLacy, J. D. Joannopoulos, M. Soljačić, *Nat. Commun.* **2014**, 5, 1.
- [42] K. Blankenbach, in *2018 IEEE Int. Conf. on Consumer Electronics (ICCE)*, IEEE, Piscataway, NJ **2018**, pp. 1–2.
- [43] Bosch, 3-d displays for use in the vehicle cockpit, https://www.bosch-mobility-solutions.com/media/global/products-and-services/passenger-cars-and-light-commercial-vehicles/human-machine-interface/visualization-systems/display-systems/bosch-3d_display_fachartikel_online.pdf (accessed: January 2021).
- [44] A. Riener, M. Rossbory, A. Ferscha, in *Workshop UX in Cars, Interact, Citeseer*, **2011**, p. 5.
- [45] F. Martinelli, F. Mercaldo, A. Orlando, V. Nardone, A. Santone, A. K. Sangaiah, *Comput. Electr. Eng.* **2020**, 83, 102504.
- [46] S. Martin, A. Tawari, E. Murphy-Chutorian, S. Y. Cheng, M. Trivedi, in *Proc. of the 4th Int. Conf. on Automotive User Interfaces and Interactive Vehicular Applications*, **2012**, pp. 149–154.
- [47] H. M. Eraqi, Y. Abouelnaga, M. H. Saad, M. N. Moustafa, *J. Adv. Transp.* **2019**, 2019, 1.

- [48] A. Firintep, M. Selim, A. Pagani, D. Stricker, in *2020 IEEE Intelligent Vehicles Symp. (IV)*, IEEE, Piscataway, NJ **2020**, pp. 1060–1065.
- [49] S. Lim, J. Jung, S.-C. Kim, S. Lee, *IEEE Access* **2020**, 8, 96606.
- [50] D. J. Yeong, G. Velasco-Hernandez, J. Barry, J. Walsh, *Sensors* **2021**, 21, 2140.
- [51] Z. Pusztai, L. Hajder, in *Proc. of the IEEE Int. Conf. on Computer Vision Workshops*, IEEE, Piscataway, NJ **2017**, pp. 394–402.
- [52] P. Moghadam, W. S. Wijesoma, D. J. Feng, in *2008 10th Int. Conf. on Control, Automation, Robotics and Vision*, IEEE, Piscataway, NJ **2008**, pp. 384–389.
- [53] P. K. Murali, I. Sorrentino, A. Rendiniello, C. Fantacci, E. Villagrossi, A. Polo, A. Ardesi, M. Maggiali, L. Natale, D. Pucci, S. Traversaro, in *Proc. of the IEEE Int. Conf. on Robotics and Automation (ICRA) 2021*, IEEE, Piscataway, NJ **2021**, pp. 1–7.
- [54] R. S. Dahiya, M. Valle, *Robotic Tactile Sensing: Technologies and System*, Springer Science & Business Media, Berlin/Heidelberg, Germany, **2012**.
- [55] G. Barrett, R. Omote, *Inform. Display* **2010**, 26, 16.
- [56] L. Skedung, M. Arvidsson, J. Y. Chung, C. M. Stafford, B. Berglund, M. W. Rutland, *Sci. Rep.* **2013**, 3, 1.
- [57] A. W. Ng, A. H. Chan, in *Proc. of the Int. Multiconf. of Engineers and Computer Scientists*, Vol. 2, **2012**, pp. 1449–1454.
- [58] R. Dahiya, N. Yogeswaran, F. Liu, L. Manjakkal, E. Burdet, V. Hayward, H. Jörintell, *Proc. IEEE* **2019**, 107, 2016.
- [59] R. Dahiya, *Proc. IEEE* **2019**, 107, 247.
- [60] R. S. Dahiya, P. Mittendorfer, M. Valle, G. Cheng, V. J. Lumelsky, *IEEE Sensors J.* **2013**, 13, 4121.
- [61] S. Khan, L. Lorenzelli, R. S. Dahiya, *IEEE Sensors J.* **2014**, 15, 3164.
- [62] R. S. Dahiya, G. Metta, M. Valle, G. Sandini, *IEEE Trans. Robotics* **2009**, 26, 1.
- [63] Z. Kappassov, J.-A. Corrales, V. Perdereau, *Robotics Auton. Syst.* **2015**, 74, 195.
- [64] A. Braun, P. Hamisu, in *Proc. of the 2nd Int. Conf. on Pervasive Technologies Related to Assistive Environments*, **2009**, pp. 1–7.
- [65] A. Braun, S. Neumann, S. Schmidt, R. Wichert, A. Kuijper, In *Proc. of the 8th Nordic Conf. on Human-Computer Interaction: Fun, Fast, Foundational*, **2014**, pp. 911–914.
- [66] A. Braun, R. Wichert, A. Kuijper, D. W. Fellner, *J. Ambient Intell. Smart Environ.* **2015**, 7, 483.
- [67] N. B. Sarter, *Int. J. Aviation Psychol.* **2000**, 10, 231.
- [68] W. Chang, W. Hwang, Y. G. Ji, *Int. J. Human-Comput. Interaction* **2011**, 27, 1119.
- [69] G. Cohen-Lazry, N. Katzman, A. Borowsky, T. Oron-Gilad, *Transp. Res. Part F: Traffic Psychol. Behav.* **2019**, 65, 217.
- [70] O. Ozioko, W. Navaraj, M. Hersh, R. Dahiya, *Sensors* **2020**, 20, 4780.
- [71] Y. Kumaresan, O. Ozioko, R. Dahiya, *IEEE Sensors J.* **2021**.
- [72] M. Bhattacharjee, M. Soni, P. Escobedo, R. Dahiya, *Adv. Electron. Mater.* **2020**, 6, 2000445.
- [73] *Immersion, Large screen touch displays, Technical report*, **2021**, https://www.immersion.com/use_cases/large-touch-screens/ (accessed: January 2021).
- [74] *Immersion: Accelerator pedals*, https://www.immersion.com/use_cases/accelerator-pedals/ (accessed: January 2021).
- [75] M. Kauer, M. Schreiber, R. Bruder, in *2010 IEEE Intelligent Vehicles Symp.*, IEEE, Piscataway, NJ **2010**, pp. 1214–1221.
- [76] M. Hirokawa, N. Uesugi, S. Furugori, T. Kitagawa, K. Suzuki, *IEEE Trans. Haptics* **2014**, 7, 334.
- [77] D. I. Katzourakis, E. Velenis, E. Holweg, R. Happee, *IEEE Trans. Intell. Transp. Syst.* **2014**, 15, 1781.
- [78] D. Blomeyer, A.-L. Schulte-Gehrmann, *ATZ Worldwide* **2019**, 121, 48.
- [79] *Yanfeng automotive interiors: Concept cars*, <https://www.yfai.com/en/concept-cars> (accessed: January 2021).
- [80] *Bmw shytech*, <https://www.bmw.de/de/neufahrzeuge/bmw-i/bmw-ix/2020/bmw-ix-ueberblick.html> (accessed: January 2021).
- [81] Y. Ma, **2020**.
- [82] J. Meyer, B. Arnrich, J. Schumm, G. Troster, *IEEE Sens. J.* **2010**, 10, 1391.
- [83] T. Hughes-Riley, C. Oliveira, R. H. Morris, T. Dias, *Digital Med.* **2019**.
- [84] C. C. Vu, J. Kim, *Sensors* **2020**, 20, 2383.
- [85] S. Takamatsu, T. Imai, T. Yamashita, T. Kobayashi, K. Miyake, T. Itoh, in *SENSORS, 2017*, IEEE, Piscataway, NJ **2017**, pp. 659–662.
- [86] *Xsensor*, <https://www.xsensor.com/> (accessed: January 2021).
- [87] *Tekscan*, <https://www.tekscan.com/> (accessed: January 2021).
- [88] M. Donneaud, C. Honnet, P. Strohmeier, in *NIME*, **2017**, pp. 7–12.
- [89] K. Qi, Y. Zhou, K. Ou, Y. Dai, X. You, H. Wang, J. He, X. Qin, R. Wang, *Carbon* **2020**, 170, 464.
- [90] F. Nikbakhtnasrabadi, H. El Matbouly, M. Ntagios, R. Dahiya, *ACS Appl. Electron. Mater.* **2021**, 3, 2233.
- [91] G. Min, Y. Xu, P. Cochran, N. Gadegaard, D. M. Mulvihill, R. Dahiya, *Nano Energy* **2021**, 83, 105829.
- [92] Y. Xu, G. Min, N. Gadegaard, R. Dahiya, D. M. Mulvihill, *Nano Energy* **2020**, 76, 105067.
- [93] L. Manjakkal, F. F. Franco, A. Pullanchiyodan, M. González-Jiménez, R. Dahiya, *Adv. Sustainable Syst.* **2021**, 5, 2000286.
- [94] R. Mukherjee, P. Ganguly, R. Dahiya, *Adv. Intell. Syst.* **2021**.
- [95] *Harman redefines the in-car concert experience with new levels of immersion and customization*, <https://news.harman.com/releases/harman-redefines-the-in-car-concert-experience-with-new-levels-of-immersion-and-customization> (accessed: January 2021).
- [96] S. J. Elliott, M. Jones, *J. Acoust. Soc. Am.* **2006**, 119, 2702.
- [97] W. Druyvesteyn, J. Garas, *J. Audio Eng. Soc.* **1997**, 45, 685.
- [98] *Hyundai showcases next-generation separated sound zone technology*, **2018**, <https://www.hyundai.news/eu/articles/press-releases/hyundai-motor-company-showcases-next-generation-separated-sound-zone-technology.html> (accessed: January 2021).
- [99] X. Liao, J. Cheer, S. Elliott, S. Zheng, *J. Audio Eng. Soc.* **2017**, 65, 226.
- [100] F. Meucci, L. Pierucci, E. Del Re, L. Lastrucci, P. Desii, in *2008 16th European Signal Processing Conf.*, IEEE, Piscataway, NJ **2008**, pp. 1–5.
- [101] B. Fazenda, H. Atmoko, F. Gu, L. Guan, A. Ball, in *2009 ICCAS-SICE*, IEEE, Piscataway, NJ **2009**, pp. 4250–4255.
- [102] V.-T. Tran, W.-H. Tsai, *IEEE Access* **2020**, 8, 75702.
- [103] L. Marchegiani, P. Newman, in *Annual Conf. Towards Autonomous Robotic Systems*, Springer, Springer Nature, Cham Switzerland, **2018**, pp. 247–259.
- [104] D. R. Seshadri, R. T. Li, J. E. Voos, J. R. Rowbottom, C. M. Alfes, C. A. Zorman, C. K. Drummond, *NPJ Digital Med.* **2019**, 2, 1.
- [105] P. Artaud, S. Planque, C. Lavergne, H. Cara, C. Tarriere, B. Gueguen, et al., in *Proc.: Int. Technical Conf. on the Enhanced Safety of Vehicles*, Vol. 1995, National Highway Traffic Safety Administration, **1995**, pp. 350–359.
- [106] A. Campagne, T. Pebayle, A. Muzet, *Physiol. Behav.* **2004**, 80, 515.
- [107] S. Sanei, J. A. Chambers, *EEG Signal Processing*, John Wiley & Sons, Inc., Hoboken, NJ, USA, **2013**.
- [108] H. Häkkinen, H. Summala, M. Partinen, M. Tiihonen, J. Silvo, *Sleep* **1999**, 22, 798.
- [109] T. C. Chieh, M. M. Mustafa, A. Hussain, S. F. Hendi, B. Y. Majlis, in *2005 1st Int. Conf. on Computers, Communications, & Signal Processing with Special Track on Biomedical Engineering*, IEEE, Piscataway, NJ **2005**, pp. 165–168.
- [110] C. Katsis, N. Ntouvas, C. Bafas, D. Fotiadis, in *Proc. of the IASTED Int. Conf. on Biomedical Engineering*, Vol. 262, **2004**.
- [111] N. Bogunovic, A. Jovic, in *17th Int. Conf. on Systems, Signals and Image Processing, IWSSIP*, **2010**, pp. 276–279.

- [112] M. Bhattacharjee, F. Nikbakhtnasrabadi, R. Dahiya, *IEEE Internet Things J.* **2021**, 8, 5101.
- [113] P. Escobedo, M. Bhattacharjee, F. Nikbakhtnasrabadi, R. Dahiya, *IEEE Internet Things J.* **2020**, 8, 5093.
- [114] M. Soni, M. Bhattacharjee, M. Ntagios, R. Dahiya, *IEEE Sens. J.* **2020**, 20, 7525.
- [115] R. Siegrwart, I. R. Nourbakhsh, D. Scaramuzza, *Introduction to Autonomous Mobile Robots*, MIT Press, Cambridge Massachusetts, USA, **2011**.
- [116] M. Jeon, M. FakhrHosseini, E. Vasey, M. A. Nees, in *Proc. of the 9th Int. Conf. on Automotive User Interfaces and Interactive Vehicular Applications Adjunct*, **2017**, pp. 178–182.
- [117] K. Harrington, D. R. Large, G. Burnett, O. Georgiou, in *Proc. of the 10th Int. Conf. on Automotive User Interfaces and Interactive Vehicular Applications*, **2018**, pp. 11–20.
- [118] J. Sterkenburg, S. Landry, M. Jeon, *J. Multimodal User Interfaces* **2019**, 13, 55.
- [119] K. G. Tippey, E. Sivaraj, W.-J. Ardoin, T. Roady, T. K. Ferris, in *Proc. of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 58, SAGE Publications, Los Angeles, CA **2014**, pp. 2023–2027.
- [120] P. R. Alves, J. Gonçalves, R. J. Rossetti, E. C. Oliveira, C. Olaverri-Monreal, in *2013 IEEE Intelligent Vehicles Symp. Workshops (IV Workshops)*, IEEE, Piscataway, NJ **2013**, pp. 1–6.
- [121] D. Dey, J. Terken, in *Proc. of the 9th Int. Conf. on Automotive User Interfaces and Interactive Vehicular Applications*, **2017**, pp. 109–113.
- [122] P. K. Murali, K. Darvish, F. Mastrogiorganni, *Intell. Service Robotics* **2020**, 13, 439.
- [123] M. Torstensson, T. H. Bui, D. Lindström, C. Englund, B. Duran, in *37th Annual Swedish Symp. on Image Analysis (SSBA 2019)*, Gothenburg, Sweden, March **2019**.
- [124] T. A. Ranney, W. R. Garrott, M. J. Goodman, *Nhtsa driver distraction research: Past, present, and future*, Technical report, SAE Technical Paper, **2001**.
- [125] J. D. Lee, K. L. Young, M. A. Regan, *Driver Distr.: Theory Effects Mitigation* **2008**, 13, 31.
- [126] S. G. Klauer, T. A. Dingus, V. L. Neale, J. D. Sudweeks, D. J. Ramsey, The impact of driver inattention on near-crash/crash risk: An analysis using the 100-car naturalistic driving study data, United States. National Highway Traffic Safety Administration, **2006**.
- [127] T. A. Dingus, F. Guo, S. Lee, J. F. Antin, M. Perez, M. Buchanan-King, J. Hankey, *Proc. Natl. Acad. Sci.* **2016**, 113, 2636.
- [128] M. J. Sullman, F. Prat, D. K. Tasci, *Traffic Injury Prev.* **2015**, 16, 552.
- [129] A. Bochkovskiy, C.-Y. Wang, H.-Y. M. Liao, *arXiv preprint arXiv:2004.10934*, **2020**.
- [130] S. Ren, K. He, R. Girshick, J. Sun, *IEEE Trans. Pattern Anal. Machine Intell.* **2016**, 39, 1137.
- [131] J. Greenberg, L. Tijerina, R. Curry, B. Artz, L. Cathey, P. Grant, D. Kochhar, K. Kozak, M. Blommer, *J. Transp. Res. Board* **2003**, 1843, 1.
- [132] W. Zhang, H. Zhang, in *IOP Conf. Series: Earth and Environmental Science*, Vol. 638, IOP Publishing, **2021**, p. 012039.
- [133] O. Nakayama, T. Futami, T. Nakamura, E. R. Boer, *SAE Trans.* **1999**, 1686.
- [134] D. Tran, H. M. Do, W. Sheng, H. Bai, G. Chowdhary, *IET Intell. Transp. Syst.* **2018**, 12, 1210.
- [135] C. Craye, A. Rashwan, M. S. Kamel, F. Karray, *Int. J. Intell. Transp. Syst. Res.* **2016**, 14, 173.
- [136] S. Park, S. D. Mello, P. Molchanov, U. Iqbal, O. Hilliges, J. Kautz, in *Proc. of the IEEE/CVF Int. Conf. on Computer Vision*, **2019**, pp. 9368–9377.
- [137] L. Alam, M. M. Hoque, in *2019 Int. Conf. on Electrical, Computer and Communication Engineering (ECCE)*, IEEE, Piscataway, NJ **2019**, pp. 1–6.
- [138] A. U. Nambi, S. Bannur, I. Mehta, H. Kalra, A. Virmani, V. N. Padmanabhan, R. Bhandari, B. Raman, in *Proc. of the 24th Annual Int. Conf. on Mobile Computing and Networking*, **2018**, pp. 840–842.
- [139] C. Ou, Q. Zhao, F. Karray, A. El Khatib, in *Int. Conf. on Image Analysis and Recognition*, Springer, **2019**, pp. 199–207.
- [140] M. V. Yeo, X. Li, K. Shen, E. P. Wilder-Smith, *Safety Sci.* **2009**, 47, 115.
- [141] T. Hwang, M. Kim, S. Hong, K. S. Park, in *2016 38th Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society (EMBC)*, IEEE, Piscataway, NJ **2016**, pp. 4646–4649.
- [142] L. R. Hartley, P. K. Arnold, G. Smythe, J. Hansen, *Appl. Ergonom.* **1994**, 25, 143.
- [143] G. F. Wilson, R. D. O'Donnell, in *Advances in Psychology*, Vol. 52, Elsevier, Amsterdam **1988**, pp. 63–100.
- [144] B. Thorslund, *Electrooculogram Analysis and Development of a System for Defining Stages of Drowsiness*, Statens väg-och transportforskningsinstitut, **2004**.
- [145] R. Fu, H. Wang, *Int. J. Neural Syst.* **2014**, 24, 1450006.
- [146] V. Balasubramanian, K. Adalarasu, *J. Bodywork Movement Therapies* **2007**, 11, 151.
- [147] R. N. Khushaba, S. Kodagoda, S. Lal, G. Dissanayake, *IEEE Trans. Biomed. Eng.* **2010**, 58, 121.
- [148] P. E. Ekman, R. J. Davidson, *The Nature of Emotion: Fundamental Questions*, Oxford University Press, Oxford **1994**.
- [149] A. Kolli, A. Fasih, F. Al Machot, K. Kyamakyia, in *Proc. of the Joint INDS'11 & ISTET'11*, IEEE, Piscataway, NJ **2011**, pp. 1–5.
- [150] T. Moriyama, K. Abdelaziz, N. Shimomura, in *Proc. of the 21st Int. Conf. on Pattern Recognition (ICPR2012)*, IEEE, Piscataway, NJ **2012**, pp. 2898–2901.
- [151] H. Gao, A. Yüce, J.-P. Thiran, in *2014 IEEE Int. Conf. on Image Processing (ICIP)*, IEEE, Piscataway, NJ **2014**, pp. 5961–5965.
- [152] Z. Ma, M. Mahmoud, P. Robinson, E. Dias, L. Skrypchuk, in *Int. Conf. on Computational Science and Its Applications*, Springer, Springer Nature, Cham Switzerland, **2017**, pp. 678–691.
- [153] A. C. Cruz, A. Rinaldi, in *Int. Conf. on Universal Access in Human-Computer Interaction*, Springer, Springer Nature, Cham Switzerland, **2017**, pp. 313–323.
- [154] M. Paschero, G. Del Vescovo, L. Benucci, A. Rizzi, M. Santello, G. Fabbri, F. F. Mascioli, in *2012 IEEE Int. Symp. on Industrial Electronics*, IEEE, Piscataway, NJ **2012**, pp. 1690–1695.
- [155] G. Rebolledo-Mendez, A. Reyes, S. Paszkowicz, M. C. Domingo, L. Skrypchuk, *IEEE Trans. Intell. Transp. Syst.* **2014**, 15, 1850.
- [156] R. R. Singh, S. Conjeti, R. Banerjee, in *2012 National Conf. on Communications (NCC)*, IEEE, Piscataway, NJ **2012**, pp. 1–5.
- [157] J. A. Healey, R. W. Picard, *IEEE Trans. Intell. Transp. Syst.* **2005**, 6, 156.
- [158] J.-S. Wang, C.-W. Lin, Y.-T. C. Yang, *Neurocomputing* **2013**, 116, 136.
- [159] R. R. Singh, S. Conjeti, R. Banerjee, *Biomed. Signal Proc. Control* **2013**, 8, 740.
- [160] S. Hoch, F. Althoff, G. McGlaun, G. Rigoll, in *Proc. (ICASSP'05), IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, 2005, Vol. 2, IEEE, Piscataway, NJ **2005**, pp. ii–1085.
- [161] M. Oehl, F. W. Siebert, T.-K. Tews, R. Höger, H.-R. Pfister, in *Int. Conf. on Human-Computer Interaction*, Springer, **2011**, pp. 577–585.
- [162] R. Taib, J. Tederry, B. Itzstein, in *CHI'14 Extended Abstracts on Human Factors in Computing Systems*, **2014**, pp. 1777–1782.
- [163] N. Mansfield, G. Sammonds, N. Darwazeh, S. Massoud, A. Mocio, T. Patel, A. Sehdev, *Proceedings of the 1st International Comfort Congress ICC* **2017**, 1.

- [164] G. M. Sammonds, M. Fray, N. J. Mansfield, *Appl. Ergonom.* **2017**, *58*, 119.
- [165] Z. Guo, H. Liu, Q. Wang, J. Yang, in *2006 8th Int. Conf. on Signal Processing*, Vol. 3, IEEE, Piscataway, NJ **2006**.
- [166] J. Wu, M. M. Trivedi, *Pattern Recogn.* **2008**, *41*, 1138.
- [167] A. Tawari, S. Martin, M. M. Trivedi, *IEEE Trans. Intell. Transp. Syst.* **2014**, *15*, 818.
- [168] A. Kondyli, V. P. Sisiopiku, L. Zhao, A. Barmoutis, *IEEE Intell. Transp. Syst. Mag.* **2015**, *7*, 18.
- [169] D. Demirdjian, C. Varri, in *2009 IEEE Workshop on Computational Intelligence in Vehicles and Vehicular Systems*, IEEE, Piscataway, NJ **2009**, pp. 16–22.
- [170] G. Andreoni, G. C. Santambrogio, M. Rabuffetti, A. Pedotti, *Appl. Ergonom.* **2002**, *33*, 511.
- [171] Z. Gao, Z. Xiao, H. Li, *J. Mech. Eng.* **2009**, *42*, 216.
- [172] A. Vergnano, F. Leali, *Human-Intell. Syst. Integration* **2020**, *2*, 35.
- [173] N. Ziraeknejad, P. D. Lawrence, D. P. Romilly, *IEEE Trans. Vehic. Technol.* **2014**, *64*, 2274.
- [174] S. A. Pullano, A. S. Fiorillo, A. La Gatta, F. Lamonaca, D. L. Carni, in *2016 IEEE International Symposium on Medical Measurements and Applications (MeMeA)*, IEEE, Piscataway, NJ **2016**, pp. 1–5.
- [175] *Bmw is about to get a personality with the company's intelligent personal assistant*, <https://bit.ly/3o3LMPv> (accessed: January 2021).
- [176] *Mbox voice assistant*, <https://www.daimler.com/magazine/technology-innovation/mbox-voice-assistant-hey-mercedes.html> (accessed: January 2021).
- [177] Y. Zheng, Y. Liu, J. H. Hansen, in *2017 IEEE Intelligent Vehicles Symp. (IV)*, IEEE, Piscataway, NJ **2017**, pp. 559–564.
- [178] M. R. Rosekind, P. H. Gander, K. B. Gregory, R. M. Smith, D. L. Miller, R. Oyung, L. L. Webbon, J. M. Johnson, *Hospital Topics* **1997**, *75*, 23.
- [179] D. R. Large, G. Burnett, V. Antrobus, L. Skrypchuk, *IET Intell. Transp. Syst.* **2018**, *12*, 420.
- [180] C. Neubauer, G. Matthews, D. Saxby, in *Proc. of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 58, SAGE Publications, Los Angeles, CA **2014**, pp. 2053–2057.
- [181] V. Antrobus, D. Large, G. Burnett, C. Hare, *J. Navigation* **2019**, *72*, 3.
- [182] P. N. Wong, D. P. Brumby, H. V. Ramesh Babu, K. Kobayashi, in *Extended Abstracts of the 2019 CHI Conf. on Human Factors in Computing Systems*, **2019**, pp. 1–6.
- [183] D. R. Large, G. Burnett, L. Clark, in *Proc. of the 11th Int. Conf. on Automotive User Interfaces and Interactive Vehicular Applications: Adjunct Proceedings*, **2019**, pp. 335–340.
- [184] K. Mahajan, D. R. Large, G. Burnett, N. R. Velaga, *Transp. Res. Part F: Traffic Psychol. Behav.* **2021**, *80*, 104.
- [185] I. Politis, P. Langdon, D. Adebayo, M. Bradley, P. J. Clarkson, L. Skrypchuk, A. Mouzakitis, A. Eriksson, J. W. Brown, K. Revell, N. Stanton, in *23rd Int. Conf. on Intelligent User Interfaces*, **2018**, pp. 601–606.
- [186] K. Abbott, in *Voice Enabling Web Applications: VoiceXML and Beyond*, Springer, Springer Nature, Cham Switzerland, **2002**, pp. 87–103.
- [187] N. Yankelovich, G.-A. Levow, M. Marx, in *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems*, **1995**, pp. 369–376.
- [188] G. M. Begany, N. Sa, X. Yuan, *Interacting Comput.* **2016**, *28*, 170.
- [189] E. Corbett, A. Weber, in *Proc. of the 18th Int. Conf. on Human-Computer Interaction with Mobile Devices and Services*, **2016**, pp. 72–82.
- [190] M. Zajicek, R. Wales, A. Lee, *Univ. Access Inform. Soc.* **2004**, *3*, 122.
- [191] J. Kim, J. H. Ryu, T. M. Han, *Etri J.* **2015**, *37*, 793.
- [192] J. L. Gabbard, G. M. Fitch, H. Kim, *Proc. IEEE* **2014**, *102*, 124.
- [193] T. Tadaki, A. Moriya, A. Hotta, T. Sasaki, H. Okumura, in *2012 IEEE Int. Symp. on Mixed and Augmented Reality (ISMAR)*, IEEE, Piscataway, NJ **2012**, pp. 323–324.
- [194] J. Jiang, Z. Huang, W. Qian, Y. Zhang, Y. Liu, *IEEE Access* **2019**, *7*, 53566.
- [195] J. L. Gabbard, J. E. Swan, D. Hix, S.-J. Kim, G. Fitch, in *2007 IEEE Virtual Reality Conf.*, IEEE, Piscataway, NJ **2007**, pp. 35–42.
- [196] N. Broy, M. Guo, S. Schneegass, B. Pfleging, F. Alt, in *Proc. of the 7th Int. Conf. on Automotive User Interfaces and Interactive Vehicular Applications*, **2015**, pp. 179–186.
- [197] Y. Gaffary, A. Lécuyer, *Frontiers in ICT* **2018**, *5*, 5.
- [198] J. Bohg, A. Morales, T. Asfour, D. Kragic, *IEEE Trans. Robotics* **2013**, *30*, 289.
- [199] J. B. Van Erp, H. Van Veen, in *Proc. of Eurohaptics*, Vol. **2001**, Eurohaptics Society, Paris, France **2001**, pp. 99–104.
- [200] S. Goldin-Meadow, *Trends Cogn. Sci.* **1999**, *3*, 419.
- [201] A. Bellarbi, S. Benbelkacem, N. Zenati-Henda, M. Belhocine, in *2011 IEEE 7th Int. Symp. on Intelligent Signal Processing*, IEEE, Piscataway, NJ **2011**, pp. 1–6.
- [202] R. Ecker, V. Broy, A. Butz, A. De Luca, in *Proc. of the 11th Int. Conf. on Human-Computer Interaction with Mobile Devices and Services*, **2009**, pp. 1–10.
- [203] M. Bhuiyan, R. Picking, in *Proceedings of the Fifth Collaborative Research Symp. on Security, E-learning, Internet and Networking (SEIN 2009)*, Citeseer, Darmstadt, Germany **2009**, pp. 25–29.
- [204] K. Kalgaonkar, B. Raj, in *2009 IEEE Int. Conf. on Acoustics, Speech and Signal Processing*, IEEE, Piscataway, NJ **2009**, pp. 1889–1892.
- [205] B. Navarro, P. K. Murali, A. Fonte, P. Fraisse, G. Poisson, A. Cherubini, in *IROS: Intelligent Robots and Systems*, IEEE/RSJ, Piscataway, NJ **2015**.
- [206] Q. Li, O. Kroemer, Z. Su, F. F. Veiga, M. Kaboli, H. J. Ritter, *IEEE Trans. Robotics* **2020**, *36*, 1619.
- [207] M. Kaboli, G. Cheng, *IEEE Trans. Robotics* **2018**, *34*, 985.
- [208] M. Kaboli, P. Mittendorfer, V. Hügel, G. Cheng, in *2014 IEEE-RAS Int. Conf. on Humanoid Robots*, IEEE, Piscataway, NJ **2014**, pp. 187–192.
- [209] M. Kaboli, D. L. R. T. Armando, R. Walker, G. Cheng, in *2015 IEEE-RAS 15th Int. Conf. on Humanoid Robots (Humanoids)*, IEEE, Piscataway, NJ **2015**, pp. 1155–1160.
- [210] B. Hjorth, *Electroencephalogr. Clin. Neurophysiol.* **1970**, *29*, 306.
- [211] M. Kaboli, A. Long, G. Cheng, *Adv. Robotics* **2015**, *29*, 1411.
- [212] M. Kaboli, G. Cheng, in *33rd Annual Conf. of the Robotics Society of Japan*, **2015**, pp. 1155–1160.
- [213] M. Kaboli, A. De La Rosa, R. Walker, G. Cheng, in *2015 IEEE-RAS 15th Int. Conf. on Humanoid Robots (Humanoids)*, **2015**, pp. 1155–1160.
- [214] M. Kaboli, K. Yao, D. Feng, G. Cheng, *Autonom. Robots* **2019**, *43*, 123.
- [215] M. Kaboli, D. Feng, K. Yao, P. Lanillos, G. Cheng, *IEEE Robot. Autom. Lett.* **2017**, *2*, 2143.
- [216] L. Graichen, M. Graichen, J. F. Krems, *Human Factors* **2019**, *61*, 774.
- [217] D. I. Katourakis, J. C. de Winter, M. Alirezaei, M. Corno, R. Happee, *IEEE Trans. Syst. Man, Cybernetics: Syst.* **2013**, *44*, 621.
- [218] E. Adell, A. Várhelyi, M. Hjälm Dahl, *Transp. Res. Part F: Traffic Psychol. Behav.* **2008**, *11*, 445.
- [219] A. Asif, S. Boll, W. Heuten, De Gruyter Oldenbourg, **2012**, *54*, 188, <https://doi.org/10.1524/itit.2012.0681>.
- [220] J. Hwang, K. Chung, J. Hyun, J. Ryu, K. Cho, in *Information Technology Convergence, Secure and Trust Computing, and Data Management*, Springer, Springer Nature, Cham Switzerland, **2012**, pp. 47–53.
- [221] T. Nukarinen, R. Raisamo, A. Farooq, G. Evreinov, V. Surakka, in *Proc. of the 8th Nordic Conf. on Human-Computer Interaction: Fun, Fast, Foundational*, **2014**, pp. 61–64.
- [222] R. Mohebbi, R. Gray, H. Z. Tan, *Human Factors* **2009**, *51*, 102.

- [223] E. C. Haas, J. B. Van Erp, *Safety Sci.* **2014**, 61, 29.
- [224] B. Thorslund, B. Peters, N. Herbert, K. Holmqvist, B. Lidestam, A. Black, B. Lyxell, *J. Eye Movement Res.* **2013**, 6, 5.
- [225] P. A. Hancock, J. E. Mercado, J. Merlo, J. B. Van Erp, *Ergonomics* **2013**, 56, 729.
- [226] J. A. Jacko, *Human Computer Interaction Handbook: Fundamentals, Evolving Technologies, and Emerging Applications*, CRC Press, **2012**.
- [227] R. Pieraccini, K. Dayanidhi, J. Bloom, J.-G. Dahan, M. Phillips, B. R. Goodman, K. V. Prasad, *New School Psychol. Bull.* **2003**, 1, 9.
- [228] B. Pfleging, S. Schneegass, A. Schmidt, in *Proc. of the 4th Int. Conf. on Automotive User Interfaces and Interactive Vehicular Applications*, ACM Digital Library, New York, USA, **2012**, pp. 155–162.
- [229] M. Braun, N. Broy, B. Pfleging, F. Alt, *J. Multimodal User Interfaces* **2019**, 13, 71.
- [230] J. Jung, S. Lee, J. Hong, E. Youn, G. Lee, in *Proc. of the 2020 CHI Conf. on Human Factors in Computing Systems*, **2020**, pp. 1–12.
- [231] S. H. Lee, S.-O. Yoon, *J. Comput. Des. Eng.* **2020**, 7, 700.
- [232] L. Angelini, J. Baumgartner, F. Carrino, S. Carrino, M. Caon, O. Khaled, J. Sauer, D. Lalanne, E. Mugellini, A. Sonderegger, in *Actes de la 28ième conférence francophone sur l'Interaction Homme-Machine*, **2016**, pp. 188–196.
- [233] A. R. Aftab, M. von der Beeck, M. Feld, in *Proc. of the 2020 Int. Conf. on Multimodal Interaction*, **2020**, pp. 595–603.
- [234] M. Mitrevska, M. M. Moniri, R. Neßelrath, T. Schwartz, M. Feld, Y. Körber, M. Deru, C. Müller, in *2015 Int. Conf. on Intelligent Environments*, IEEE, Piscataway, NJ **2015**, pp. 180–183.
- [235] F. Roider, S. Rümelin, B. Pfleging, T. Gross, in *Proc. of the 9th Int. Conf. on Automotive User Interfaces and Interactive Vehicular Applications*, **2017**, pp. 94–102.
- [236] R. Neßelrath, M. M. Moniri, M. Feld, in *2016 12th Int. Conf. on Intelligent Environments (IE)*, IEEE, Piscataway, NJ **2016**, pp. 190–193.
- [237] *Natural and fully multimodal interaction with the vehicle and its surroundings*, <https://bit.ly/2UagYS9> (accessed: January 2021).
- [238] K. M. Bah, M. G. Jæger, M. B. Skov, N. G. Thomassen, in *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems*, **2008**, pp. 1139–1148.
- [239] T. Kopinski, J. Eberwein, S. Geisler, U. Handmann, in *2016 IEEE 19th Int. Conf. on Intelligent Transportation Systems (ITSC)*, IEEE, Piscataway, NJ **2016**, pp. 661–666.
- [240] *Road Vehicles – ergonomic aspects of transport information and control systems – simulated lane change test to assess driver distraction*, <https://www.iso.org/obp/ui/#iso:std:iso:26022:ed-1:v1:en> (accessed: January 2021).
- [241] S. Rümelin, C. Marouane, A. Butz, in *Proc. of the 5th Int. Conf. on Automotive User Interfaces and Interactive Vehicular Applications*, **2013**, pp. 40–47.
- [242] A. B. Koyuncu, E. Erçelik, E. Comulada-Simpson, J. Venrooij, M. Kaboli, A. Knoll, in *2020 IEEE Intelligent Vehicles Symp. (IV)*, IEEE, Piscataway, NJ **2020**, pp. 2118–2125.
- [243] S. Li, W. Deng, *IEEE Trans. Affect. Comput.* **2020**.
- [244] C. Shan, S. Gong, P. W. McOwan, *Image Vision Comput.* **2009**, 27, 803.
- [245] J. Zeng, S. Shan, X. Chen, in *Proc. of the European Conf. on Computer Vision (ECCV)*, **2018**, pp. 222–237.
- [246] X. Chen, Q. Lin, D. Zhou, in *Int. Conf. on Machine Learning. PMLR*, **2013**, pp. 64–72.
- [247] Y. Zhang, X. Chen, D. Zhou, M. I. Jordan, *J. Machine Learn. Res.* **2016**, 17, 3537.
- [248] H. Yang, J. Shi, L. Carlone, *IEEE Trans. Robotics* **2020**.
- [249] M. Kaboli, *A Review of Transfer Learning Algorithms*, Research report, Technische Universität München, Transfer Learning Algorithms, **2017**, <https://hal.archives-ouvertes.fr/hal-01575126>.
- [250] P. Falco, S. Lu, C. Natale, S. Pirozzi, D. Lee, *IEEE Trans. Robotics* **2019**, 35, 987.
- [251] T. Tommasi, F. Orabona, M. Kaboli, B. Caputo, *Br. Machine Vision (BMVC)* **2012**.
- [252] M. Kaboli, R. Walker, G. Cheng, in *2016 IEEE Int. Conf. on Robotics and Automation (ICRA)*, **2016**, pp. 2242–2247.
- [253] D. Feng, M. Kaboli, G. Cheng, *Sensors* **2018**, 18, 2.
- [254] M. Kaboli, D. Feng, G. Cheng, *Int. J. Humanoid Robotics* **2018**, 15, 1850001.
- [255] P. K. Murali, M. Gentner, M. Kaboli, in *Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)* **2021**, **2021**, pp. 1–7.
- [256] K.-H. Glander, *ATZelectron. Worldwide* **2020**, 15, 66.
- [257] L. Guo, S. Mangani, Y. Liu, Y. Jia, *IEEE Trans. Vehic. Technol.* **2018**, 67, 8085.
- [258] S. Cao, P. Tang, X. Sun, *Safety* **2020**, 6, 34.
- [259] M. Hartstern, V. Rack, M. Kaboli, W. Stork, in *2020 IEEE Intelligent Vehicles Symp. (IV)*, IEEE, Piscataway, NJ **2020**, pp. 858–864.
- [260] M. Ruff, in *2003 IEEE 58th Vehicular Technology Conf. VTC 2003-Fall (IEEE Cat. No. 03CH37484)*, Vol. 5, IEEE, Piscataway, NJ **2003**, pp. 3382–3389.
- [261] I. ISO 1898.
- [262] R. Makowitz, C. Temple, in *2006 IEEE Int. Workshop on Factory Communication Systems*, IEEE, Piscataway, NJ **2006**, pp. 207–212.
- [263] F. Hartwich, et al., *Proceedings of the 13th international CAN Conference (ICC 2012)*, Citeseer, **2012**, pp. 1–9.
- [264] B. Kraemer, *IEEE Commun. Mag.* **2016**, 54, 4.
- [265] K. Matheus, T. Königseder, *Automotive Ethernet*, Cambridge University Press, Cambridge UK, **2021**.
- [266] H. Ullah, N. G. Nair, A. Moore, C. Nugent, P. Muschamp, M. Cuevas, *IEEE Access* **2019**, 7, 37251.
- [267] P. T. Internet, *Application scenarios, definitions and terminology, architecture, functions, and technical assumptions*, https://standards.ieee.org/project/1918_1.html (accessed: January 2021).
- [268] D. Jia, K. Lu, J. Wang, X. Zhang, X. Shen, *IEEE Commun. Surv. Tutorials* **2015**, 18, 263.
- [269] F. Dressler, F. Klingler, M. Segata, R. L. Cigno, *Proc. IEEE* **2018**, 107, 436.
- [270] A. Aijaz, M. Sooriyabandara, *Proc. IEEE* **2018**, 107, 414.
- [271] J. S. Warner, R. G. Johnston, *J. Sec. Adminis.* **2002**, 25, 19.
- [272] T. E. Humphreys, B. M. Ledvina, M. L. Psiaki, B. W. O'Hanlon, P. M. Kintner, et al., in *Proc. of the 21st Int. Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS 2008)*, **2008**, pp. 2314–2325.
- [273] N. O. Tippenhauer, C. Pöpper, K. B. Rasmussen, S. Capkun, in *Proc. of the 18th ACM Conf. on Computer and Communications Security*, **2011**, pp. 75–86.
- [274] H. Shin, D. Kim, Y. Kwon, Y. Kim, in *Int. Conf. on Cryptographic Hardware and Embedded Systems*, Springer, **2017**, pp. 445–467.
- [275] J. Petit, B. Stottelaar, M. Feiri, F. Kargl, *Black Hat Europe* **2015**, 11, 995.
- [276] K. Ren, Q. Wang, C. Wang, Z. Qin, X. Lin, *Proc. IEEE* **2019**, 108, 357.
- [277] C. Miller, C. Valasek, *Black Hat USA* **2015**, 2015, S91.
- [278] S. C. HPL, *Application Report SLOA101*, **2002**, pp. 1–17.
- [279] K.-T. Cho, K. G. Shin, in *Proc. of the 2016 ACM SIGSAC Conf. on Computer and Communications Security*, **2016**, pp. 1044–1055.
- [280] N. Khatri, R. Shrestha, S. Y. Nam, *Electronics* **2021**, 10, 893.
- [281] G. Goh, N. Cammarata, C. Voss, S. Carter, M. Petrov, L. Schubert, A. Radford, C. Olah, *Distillation* **2021**, 6, e30.
- [282] C. Sitawarin, A. N. Bhagoji, A. Mosenia, M. Chiang, P. Mittal, in *Proc. of ACM Conf. on Computer and Communications Security*, **2018**.
- [283] Q. Zhang, H. Zhong, J. Cui, L. Ren, W. Shi, *IEEE Internet Things J.* **2020**.
- [284] H. Li, D. Ma, B. Medjahed, Y. S. Kim, P. Mitra, *SAE Int. J. Adv. Curr. Pract. Mobility* **2019**, 1, 1035.

- [285] A. Rendiniello, A. Remus, I. Sorrentino, P. Murali, P. Kumar, M. Daniele, N. Marco, T. Lorenzo, V. Silvio, P. Enrico, A. A. Andrea, *IEEE Internet Things J.* **2020**.
- [286] U. Neisser, R. Becklen, *Cognitive Psychol.* **1975**, 7, 480.
- [287] M. Yeh, C. D. Wickens, *Attention and Trust Biases in the Design of Augmented Reality Displays*, Technical report, Illinois University at Urbana-Champaign Savoyaviation Research Lab, **2000**.
- [288] Data is the new oil in the future of automated driving, <https://newsroom.intel.com/editorials/krzanich-the-future-of-automated-driving/#gs.tutt49> (accessed: January 2021).
- [289] J. Tang, S. Liu, L. Liu, B. Yu, W. Shi, *IEEE Access* **2020**, 8, 30467.
- [290] R. Bajcsy, *Proc. IEEE* **1988**, 76, 966.
- [291] J. A. Baxter, D. A. Merced, D. J. Costinett, L. M. Tolbert, B. Ozpineci, in *2018 IEEE Transportation Electrification Conf. and Expo (ITEC)*, IEEE, Piscataway, NJ **2018**, pp. 944–949.
- [292] Dds foundation, <https://www.dds-foundation.org/> (accessed: January 2021).
- [293] M. Quigley, K. Conley, B. Gerkey, J. Faust, T. Foote, J. Leibs, R. Wheeler, A. Y. Ng, et al., in *ICRA Workshop on Open Source Software*, Vol. 3, Kobe, Japan **2009**, p. 1.
- [294] Autosar, <https://www.autosar.org/> (accessed: January 2021).
- [295] A. Rendiniello, A. Remus, I. Sorrentino, P. K. Murali, D. Pucci, M. Maggiali, L. Natale, S. Traversaro, E. Villagrossi, A. Polo, et al., in *2020 25th IEEE Int. Conf. on Emerging Technologies and Factory Automation (ETFA)*, Vol. 1, IEEE, Piscataway, NJ **2020**, pp. 1273–1276.
- [296] J. Lee, K. M. Kockelman, in *98th Annual Meeting of the Transportation Research Board*, Vol. 1, Washington, DC **2019**.
- [297] M. H. Mobarak, R. N. Kleiman, J. Bauman, *IEEE Trans. Transp. Electrification* **2020**.
- [298] P. Escobedo, M. Ntagios, D. Shakthivel, W. T. Navaraj, R. Dahiya, *IEEE Trans. Robotics* **2020**.
- [299] C. B. Robledo, V. Oldenbroek, F. Abbruzzese, A. J. van Wijk, *Appl. Energy* **2018**, 215, 615.
- [300] Self-driving cars use crazy amounts of power, and it's becoming a problem, <https://www.wired.com/story/self-driving-cars-power-consumption-nvidia-chip/> (accessed: January 2021).
- [301] A. Labs, *Driving intelligence at the edge with neuromorphic computing*, https://www.accenture.com/_acnmedia/PDF-145/Accenture-Neuromorphic-Computing-POV.pdf#zoom=40 (accessed: January 2021).
- [302] D. Marković, A. Mizrahi, D. Querlioz, J. Grollier, *Nat. Rev. Phys.* **2020**, 2, 499.
- [303] M. Soni, R. Dahiya, *Philos. Trans. R. Soc. A* **2020**, 378, 20190156.
- [304] W. Taube Navaraj, C. Garca Núñez, D. Shakthivel, V. Vinciguerra, F. Labeau, D. H. Gregory, R. Dahiya, *Front. Neurosci.*, *Frontiers*, **2017**, 11, 501.
- [305] N. Yogeswaran, W. Dang, W. T. Navaraj, D. Shakthivel, S. Khan, E. O. Polat, S. Gupta, H. Heidari, M. Kaboli, L. Lorenzelli, et al., *Adv. Robotics* **2015**, 29, 1359.
- [306] The mercedes-benz f 015 luxury in motion, <https://www.mercedes-benz.com/en/innovation/autonomous/research-vehicle-f-015-luxury-in-motion/> (accessed: January 2021).
- [307] The bmw group at the consumer electronics show (ces) 2020 in las vegas, <https://www.press.bmwgroup.com/global/article/detail/T0304187EN/the-bmw-group-at-the-consumer-electronics-show-ces-2020-in-las-vegas?language=en> (accessed: January 2021).
- [308] W. Zeng, M. A. Khalid, S. Chowdhury, *IEEE Commun. Surv. Tutorials* **2016**, 18, 1552.
- [309] K. French, *Energy Consumption of In-Vehicle Communication in Electric Vehicles: A Comparison between can, Ethernet and IEEE 802.11p*, **2019**.
- [310] Tja1020: Lin transceiver, <https://www.nxp.com/products/interfaces/automotive-lin-solutions/lin-transceiver:TJA1020> (accessed: January 2021).
- [311] Tja1050 high speed can transceiver, <https://www.nxp.com/docs/en/data-sheet/TJA1050.pdf> (accessed: January 2021).
- [312] Tja1100 100base-t1 phy for automotive ethernet, <https://www.nxp.com/docs/en/data-sheet/TJA1100.pdf> (accessed: January 2021).
- [313] Tja1080ats: Flexray transceiver, <https://www.nxp.com/products/interfaces/flexray-transceivers/flexray-transceiver:TJA1080ATS> (accessed: January 2021).



Prajval Kumar Murali is a Ph.D. candidate with the BMW Group and the Bendable Electronics and Sensing Technologies (BEST) group, University of Glasgow. Previously, he worked as a robotics engineer with the Dynamic Interaction Control group at the Istituto Italiano di Tecnologia, Genova. He received the master's degree in robotics engineering jointly from the University of Genova, Italy and Ecole Centrale Nantes, France and was awarded the Erasmus-Mundus Consortium Scholarship. His research interests include visio-tactile perception, human–robot interaction, and robot control.



Mohsen Kaboli is an assistant professor and lead of robotics, AI, and tactile intelligence at Radboud University and the BMW Group research. Previously, he was a group leader and postdoctoral research fellow at the Institute for Advanced Studies (IAS), the Technical University of Munich (TUM), Germany from September 2017 till August 2018. He received his Ph.D. degree with the highest distinction (summa cum laude) in robotics and machine learning focusing on tactile perception and learning in robotics from TUM, Germany, in 2017. He was awarded the best European Ph.D. thesis award in robotics, Georges Giralt Ph.D. Award (finalist).



Ravinder Dahiya is a professor of electronics and nanoengineering at the University of Glasgow, UK. He is the leader of Bendable Electronics and Sensing Technologies (BEST) research group, which conducts multidisciplinary research in flexible and printed electronics, electronic skin, robotics, and wearable systems. He has authored about 400 articles, 7 books, and submitted/granted patents and disclosures. He has led several international projects. He is president-elect (2020–2021) of IEEE Sensors Council and Founding Editor-in-Chief of IEEE Journal on Flexible Electronics. He has chaired several conferences including IEEE FLEPS, which he founded.