



Pioneering In-Cabin Monitoring

Unmasking the Power of 2D and 3D Cameras

Table of contents

About Fraunhofer IOSB	6
1. Why In-Cabin Monitoring?	7
1.1 Legislation and Multilateral Agreements	8
1.2 User Requirements	14
1.3 Technology providers	15
1.4 Realizing DMS and OMS Applications with a ToF sensor	16
1.5 Consumer Electronics	18
1.6 Interview with Quirin Anker from Daimler Truck (FUSO) about DMS for Trucks	19
1.7 Marco Willems from CARIAD about Digital Twins	21
2. Applications for In-Cabin Monitoring	23
2.1 Active Safety	24
2.2 Passive Safety	24
2.3 Entertainment	25
2.4 Comfort	25
2.5 AI Personal Assistants	26
3. Roadmap for In-Cabin Monitoring Applications	27
4. Cameras for In-Cabin Monitoring	28
4.1 What kind of camera sensors are suitable?	28
4.2 Interview with Elena Zhelondz from A2MAC1 about In-Cabin Sensors in Series Production Vehicles	31
4.3 Comparison of sensors	33
4.4 Interview with Prof. Dr. Rainer Stiefelhagen	34
4.5 Interview with Prof. Dr. Jürgen Beyerer	35
4.6 2D vs. 3D sensing	37
4.7 The dependency of camera position, lens opening angle and resolution	37
4.8 Evaluation of ToF for in-cabin monitoring	38
5. State of the Art in Sensors, Training Data and Algorithms	42
6. Conclusions	44
References	45

List of figures

Figure 1: Overview Legislation regarding in-cabin monitoring	7
Figure 2: Overview automotive safety institutions	8
Figure 3: Overview consumer test organizations	13
Figure 4: Metaphor for in-cabin monitoring systems as a butler	14
Figure 5: Overview of applications from user requirements study	15
Figure 6: Selection of companies working on DMS	15
Figure 7: Overview applications on market	16
Figure 8: Melexis ToF sensor specifications	16
Figure 9: OMS detection features from combination of NIR and depth image	17
Figure 10: DMS detection features with Melexis ToF and its NIR image sensor	18
Figure 11: Overview applications from consumer electronics	18
Figure 12: Quirin Anker from Daimler Truck (FUSO)	19
Figure 13: Overview of applications	23
Figure 14: Start with the prio 1 application to decide on the camera technology which will enable further applications.	28
Figure 15: Elena Zhelondz, A2MAC1	31
Figure 16: Sensors used for in-cabin monitoring	31
Figure 17: Most popular locations of DMCs	32
Figure 18: Average cost of interior sensors	32
Figure 19: Prof. Dr. Rainer Stiefelhagen	34
Figure 20: Prof. Dr. Jürgen Beyerer	35
Figure 21: Typical camera positions with different lens opening angles	38
Figure 22: Martin Lass, Infineon	39
Figure 23: Overview of the measured features for the evaluation	40
Figure 24: Simulated sensor data for five different camera positions and two different secondary activities of the driver	40
Figure 25: ChatGPT4 in April, 2024, alerting car drivers based on an in-cabin image	43
Figure 26: Drive & Act dataset for driver activity recognition training	43
Figure 27: Expected technology leaps in in-cabin monitoring software	44

List of tables

Table 1: EU legislation and binding dates of system integration	9
Table 2: Comparison of sensors	30
Table 3: Detected features of sensors	33
Table 4: Technical suitability of sensors	33
Table 5: Applications and suitable sensors	34
Table 6: Pros and Cons of 2D and 3D interior monitoring sensors	37
Table 7: Comparison of different features of the driver and co-driver for different camera positions with Infineon VGA resolution and 120° x 90° lens	41



Lead Author

Dr. Frederik Diederichs
Fraunhofer IOSB

Acknowledgements

Our thanks belong to all supporters of this paper, to interview partners, contributors, sponsors, reviewers, and sparring partners.

Namely

- Kathleen Entz
- Amina Hermanns
- Manuel Martin
- David Lerch
- Christian Lengenfelder
- Gerrit Holzbach
- Zeyun Zhong
- Jutta Hild
- Michael Voit
- Ajona Vijayakumar
- Quirin Anker
- Martin Lass
- Elena Zhelondz
- Kristof Lieben
- Thomas Parton
- Rainer Stiefelhagen
- Jürgen Beyerer

About Fraunhofer IOSB

As part of the largest organization for application-oriented research in Europe, the Fraunhofer Institute of Optronics, System Technologies, and Image Exploitation IOSB, with headquarters in Karlsruhe, is one of the leading scientific institutes in the fields of Artificial Intelligence (AI), Computer Vision and Optics in Germany and Europe. Approximately 850 employees research and support companies in optimizing products, services and processes and in developing new digital business models. Fraunhofer IOSB is shaping the digital transformation of our working and living environments: with innovative AI applications for industry, health and sustainability, with forward-looking computer vision technologies and extensive optical sensor know how.

In the department “Human-AI Interaction”, innovative interaction methods and assistance systems are developed to support people in their tasks. With the development of camera-based perception and adaptive user interfaces, the focus is particularly on the detection of humans and the evaluation of their activities for multimodal human-machine interactions in intelligent and proactive environments. For more than 10 years we have been focussing on vehicle cabins.

Numerous publications resulted from the research on occupant monitoring:

- Martin, M., Lerch, D., & Voit, M. (2023, June). Viewpoint Invariant 3D Driver Body Pose-Based Activity Recognition. In 2023 IEEE Intelligent Vehicles Symposium (IV) (pp. 1–6). IEEE.
- Martin, M., Voit, M. & Stiefelhagen, R. (2021). An Evaluation of Different Methods for 3D-Driver-Body-Pose Estimation. IEEE. <https://doi.org/10.1109/itsc48978.2021.9564676>
- Martin, M., Voit, M. & Stiefelhagen, R. (2020). Dynamic Interaction Graphs for Driver Activity Recognition. 2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC). <https://doi.org/10.1109/itsc45102.2020.9294520>
- Martin, M., Roitberg, A., Haurilet, M., Horne, M., Reib, S., Voit, M. & Stiefelhagen, R. (2019). Drive & Act: A Multi-Modal Dataset for Fine-Grained Driver Behavior Recognition in Autonomous Vehicles. Proceedings of the IEEE International Conference on Computer Vision. <https://doi.org/10.1109/iccv.2019.00289>
- Roitberg, A., Pollert, T., Haurilet, M., Martin, M. & Stiefelhagen, R. (2019). Analysis of Deep Fusion Strategies for Multi-Modal Gesture Recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. <https://doi.org/10.1109/cvprw.2019.00029>
- Ludwig, J., Martin, M., Horne, M., Flad, M., Voit, M.,

Stiefelhagen, R. & Hohmann, S. (2018). Driver observation and shared vehicle control: supporting the driver on the way back into the control loop. *Automatisierungstechnik*, 66(2), 146–159.

<https://doi.org/10.1515/auto-2017-0103>

- Martin, M., Popp, J., Anneken, M., Voit, M. & Stiefelhagen, R. (2018). Body Pose and Context Information for Driver Secondary Task Detection. IEEE. <https://doi.org/10.1109/ivs.2018.8500523>
- Martin, M., Stuehmer, S., Voit, M. & Stiefelhagen, R. (2017). Real time driver body pose estimation for novel assistance systems. 2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC). <https://doi.org/10.1109/itsc.2017.8317722>
- Martin, M., Van De Camp, F. & Stiefelhagen, R. (2014). Real Time Head Model Creation and Head Pose Estimation on Consumer Depth Cameras. IEEE. <https://doi.org/10.1109/3dv.2014.54>

The Advanced Occupant Monitoring System from Fraunhofer IOSB uses optical sensors inside the vehicle. It captures the cabin, the driver and all occupants, recognizes the 3D body pose of all individuals, analyses their movement behaviour, and classifies the activity of each detected person. With this, it is not only possible to recognize critical situations, such as a driver falling asleep, but also to distinguish between various activities and the associated attention levels. This supports both safety systems and comfort functions inside the vehicle.

Fraunhofer IOSB technology has been developed in joint projects with leading industry partners and funding from German Government. In the BMBF projects InCarIn and PAKoS, the Fraunhofer IOSB developed machine learning-based methods to capture the body pose and interactions of all vehicle occupants. With these methods, functions in the vehicle were adjusted and personalized. In the BMWK project INITIATIVE (www.initiative-projekt.de), the detection was expanded to pedestrians in traffic and applied to the interaction of pedestrians with automated vehicles. In the BMWK funded project KARLI (www.karli-projekt.de), these skills are currently being transferred to new AI methods of unsupervised learning.

<https://www.iosb.fraunhofer.de/de/kompetenzen/bildauswertung/interaktive-analyse-diagnose/automotive.html>

www.iosb.fraunhofer.de

1. Why In-Cabin Monitoring?

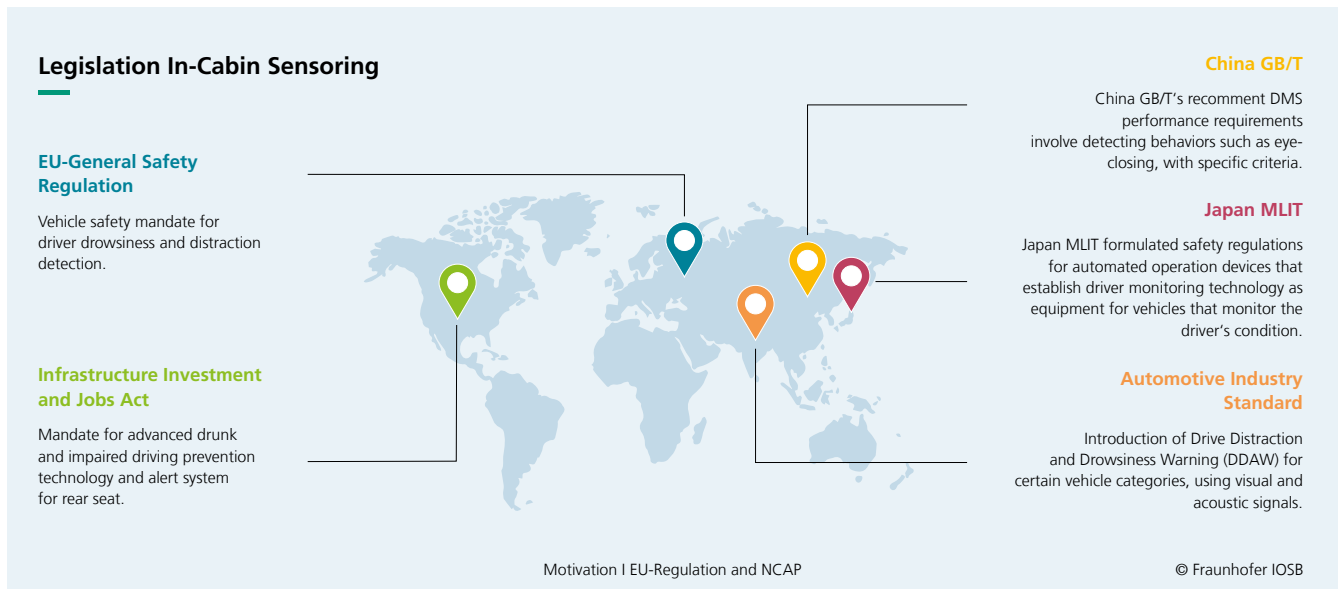


Figure 1: Overview Legislation regarding in-cabin monitoring

According to the World Health Organization, approximately 1.19 million people die each year in road traffic accidents [1]. National and international studies indicate that driver distraction and inattention are significant factors in traffic accidents [2]. In 2022, 3308 people were killed, and an estimated additional 289,310 people were injured in motor vehicle traffic crashes involving distracted drivers [3]. Automation of driving tasks stimulates further distraction, sleepiness and operating vehicles under reduced driving ability.

Regulatory authorities in the automotive industry in Europe, America, and Asia are paving the way for safety technologies.

In-cabin monitoring and driver monitoring are considered major contributors to improving traffic safety in the future. This expectation has led to the enactment of new and expanded laws, compelling automotive manufacturers to integrate in-cabin monitoring into vehicles.

Improving traffic safety is the primary driver for the introduction of in-cabin monitoring and driver monitoring. Worldwide legislation defines the causes of traffic accidents and injuries that should be mitigated with in-cabin monitoring.

Consumer test associations design and publish vehicle assessments in parallel with the ambitious demands from legislation.

These assessments define how in-cabin monitoring is tested and which key performance indicators need to be achieved.

Hence technology providers have increased their efforts to offer sensors, algorithms and systems to comply to legislation and assessments. A large variety of options is currently under development.

Beyond safety applications, in-cabin monitoring can and will be used for comfort and entertainment and will also play a fundamental role in creating personalized digital services for car users. In parallel to services on smartphones, AR headsets and in smart home environments, the vehicle interior may become one of the most digitalized and personalized environments for humans to be – thanks to in-cabin monitoring.

Hence, the development of in-cabin monitoring is currently driven by five major forces:

1. Legislation and multilateral agreements from governments
2. Vehicle assessment programs and consumer tests
3. Technology advancements
4. Consumer electronics
5. User requirements

1.1 Legislation and Multilateral Agreements

Due to the impact of human error in traffic safety, governments around the world have identified interior monitoring and driver monitoring as a key to improve safety. Furthermore, advancements in automated driving require humans as operators and fallback actors for the automated systems. Ensuring human availability and reliable performance is hence a safety relevant factor. Many legislations for the introduction of automated driving require monitoring systems to ensure the human fallback layer.

Many countries have already passed or are preparing such regulation.

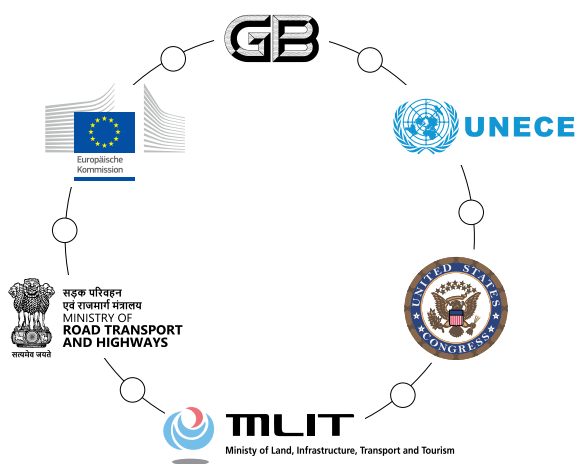


Figure 2: Overview automotive safety institutions

Europe

The European Parliament adopted a new regulation on the general safety of vehicles (GSR) and the protection of vehicle occupants and vulnerable road users in November 2019 [4]. The General Safety regulation specifies and focuses on requirements for Driver Monitoring Systems (DMS), which have a high potential to significantly reduce the number of casualties.

On November 22, 2023, the European Commission published a new Delegated Regulation (EU) 2023/2590, within necessary technical requirements and test procedures were developed and established to support ADDW systems [5]. The regulation becomes mandatory from July 7, 2024 for all passenger carrying and goods carrying motor vehicles and becomes requisite to all new vehicles of type M and N category vehicle from July 7, 2024.

National authorities are not allowed to refuse EU type-approval-storequesting manufacturers for new vehicle types or the extension of approval for an existing vehicle type starting from January 1, 2024, provided that the vehicle complies with the present regulations.

General requirements in GSR 2019/2144 [4]

Regulations and test procedures for the type of approval of vehicles and systems shall be established and harmonized at the Union level. The GSR requires besides others the installation of the following driver monitoring systems in all motor vehicles (see Table 1):

- **Driver drowsiness and attention warning (DDAW):** a system that assesses the driver's alertness through vehicle systems analysis and warns the driver if needed.
- **Advanced driver distraction warning (ADDW):** a system that helps the driver to continue to pay attention to the traffic situation and that warns the driver when he or she is distracted.

The equipment of automated vehicles includes:

- **Driver availability monitoring systems:** a system to assess whether the driver is able to take over the driving function from an automated vehicle in particular situations, where appropriate.

Specific requirements of ADDW in GSR 2023/2590 [5]

The publication specifies necessary technical requirements and test procedures for ADDW systems to monitor the direction of the driver's gaze. Warning signals are activated when the driver's gaze deviates from the specified areas within a predetermined time, thus being defined as "distracted".

Defined distraction zones

- Area 1
 - Roof
 - Left side (-55°)
 - Right side (+55°)
- Area 2
 - Area of the windshield and windows
- 10° around the area of the windshield and windows, measured from the eye reference point
- Area 3: Any surface below a plane extending 30° downward from the driver's eye reference point
- **Operating and ambient conditions**
 - ADDW system must be automatically activated at speeds above 20 km/h
 - Measurement and calibration of the system permitted with a cumulative travel time of up to 1 minute of driving at speeds ≥ 20 km/h
 - System must operate under both daytime and night-time conditions
- **Performance indicators**
 - For speeds between 20 km/h and 50 km/h the AADWS system must provide a warning if the drivers gaze remains within the defined "distracted" vision are for more than 6 seconds.
 - Remains the drivers gaze within the specified areas of distraction for more than 3.5 seconds and a speed above 50 km/h, the ADDW system must give a warning.

■ Warnings

- Is driver distraction detected, the ADDW-system must provide a warning strategy.
- Driver must be informed by visual warning, plus an acoustic and/or direct/indirect haptic warning.
- Intensification and escalation of warnings until the driver distraction ceases

■ System failure

- Issuance of an error warning signal upon detection of a temporary non-electrical fault condition as well as permanent failures.

■ Roadworthiness testing provisions

- For the purpose of regular technical monitoring of vehicles, the regulation specifies the features of the ADDW for verification.

Test procedures and technical requirements

In addition to Regulation (EU) 2019/2144 of the European Parliament, the European Commission published detailed rules on the specific test procedures and technical requirements for the type-approval of motor vehicles regarding their fatigue and alertness warning systems (see ANNEXES to the Commission Delegated Regulation supplementing Regulation (EU) 2019/2144) [6].

EU-Legislation

Driver Monitoring Systems	Dates	Explanatory Notes
Driver Drowsiness and Attention Warning (DDAW)	July 6, 2022	The GSR mandates the mandatory installation of DDAW (Driver Distraction and Attention Warning) systems in all motor vehicles for the carriage of passengers and goods (new vehicle types of classes M and N) starting from July 6, 2022. (see GSR, Annex II, E2) [4]
	July 7, 2024	DDAW (Driver Distraction and Attention Warning) systems will be mandated for all new motor vehicles for the carriage of passengers and goods starting from July 7, 2024. Vehicles, as well as the commissioning of components and autonomous technical units that do not meet the requirements, are excluded from approval. (see GSR, Annex II, Part E2) [4]
Advanced Driver Distraction Warning (ADDW)	July 7, 2024	Regulation (EU) 2019/2144 mandates the compulsory installation of an ADDW (Advanced Driver Distraction Warning) system starting from July 7, 2024, for all new vehicle types of classes M and N. (see GSR, Annex II, E3) [4]
	July 7, 2026	From July 7, 2026, onwards, ADDW (Advanced Driver Distraction Warning) systems will be mandatory for all new motor vehicles for the carriage of passengers and goods. Vehicles, as well as the commissioning of components and autonomous technical units that do not meet the requirements, are excluded from approval. (see GSR, Annex II, E3 and GSR2 2023/2590) [4] [5]
Driver Availability Monitoring System (DAMS)	July 6, 2022	Regulation (EU) 2019/2144 stipulates the mandatory installation of a DAMS system from July 6, 2022 for all new vehicle types in categories M and N. Compliance is required in case of automated vehicles. (see GSR, Annex II, E4) [4]
	July 7, 2024	From July 7, 2024, DAMS systems will be mandatory for all new motor vehicles for the transportation of passengers and goods. Vehicles, as well as the commissioning of components and separate technical units that do not meet the requirements, are excluded from registration. Compliance is required in case of automated vehicles. (see GSR, Annex II, E4) [4]

Table 1: EU legislation and binding dates of system integration

United States

The United States have introduced several regulations for driver monitoring systems (DMS).

Hot Cars Act of 2021

The Helping Overcome Trauma for Children Alone in Rear Seats, HOT Cars Act of 2021 [7], was introduced in 2021. The bill aims to make child safety alert systems mandatory in all new passenger vehicles in order to reduce death and injury resulting from heatstroke in vehicles. Within two years after the date of enactment of this section, a final rule shall be published requiring all new passenger vehicles to be equipped with a system to detect children that are left in a parked car. The Act was referred to the Subcommittee on Consumer Protection and Commerce and remains on status *Introduced*.

Safe Act of 2021

The Stay Aware for Everyone Act of 2021 (SAFE Act of 2021) mandates the use of driver monitoring systems to minimize or eliminate:

- **driver distraction**
- **driver disengagement**
- **automation complacency by drivers**
- **foreseeable misuse of advanced driver-assist systems**

A final rule establishing performance standards and the installation of driver monitoring systems is due within four years of the date at the latest. Further, automakers are instructed to comply within two model years of the effective date of the final rule. As of right now, the Act remains in Committee on Commerce, Science, and Transportation on status *Introduced* [8].

Infrastructure Investment and Jobs Act

The Infrastructure Investment and Jobs Act (IIJA), from 2021 calls in section B for the use of driver monitoring systems to be investigated and for a corresponding DMS regulation to be issued once the research has been successfully completed. The Act became Public Law by November 15, 2021 [9]. Specifically, the act demands:

- **Reduction of driver distraction (HR 3684; Section 24209/P. 1045)**

The Safe Act was included in the IIJA with its requirements for research into driver monitoring systems for use against driver distraction. By November 15, 2024, the Secretary must conduct research on the installation and use of DMS in motor vehicles to minimize or eliminate.

1. Driver distraction
2. Driver disengagement
3. Driver complacency with automation
4. predictable misuse of advanced driver assistance systems.

- **Advanced Impaired Driving Prevention Technology (HR 3684; Section 24220/P. 1066)**

Within three years of the entry into force of this Act, a standard shall be issued that requires all new passenger motor vehicles to be equipped with advanced drunk and impaired driving prevention technology as standard.

Required is a system that can passively and precisely monitor the driver's performance and blood concentration and verify for impairment. Therefore it is due on November 15, 2024 of this year.

- **Awareness of children in motor vehicles (HR 32304; Section 32304B/P. 1077)**

Directs the secretary to conduct a study on the possible retrofitting of existing passenger vehicles with one or more technologies that reduce the risk of children being left in the rear seats after a motor vehicle is disabled. Section contains parts of the HOT Cars Act.

Alliance of Automobile Manufacturers

Participating car manufacturers in the Alliance of Automobile Manufacturers have made an independent and voluntary commitment to integrate rear seat reminder systems as standard equipment in passenger cars by model year 2025 at the latest. This measure serves to protect and minimize cases in which children are left alone in the vehicle. According to the Alliance for Automotive Innovation Press Release of November 1, 2023, rear seat technology is already available in more than 215 new vehicle models (standard and optional) [10]. The rear seat reminder system uses radar to monitor and detect movement at the rear seats and alert the driver through various acoustic or visual options.

China

In 2018, Jiangsu became the first Chinese province to require the use of driver monitoring systems in long-haul trucks and vehicles transporting dangerous goods.

At the present time, the China-GB/T does not set any binding requirements, but rather makes an official recommendation on performance requirements for DMS. This recommended national standard (GB/T 41797-2022) entitled Driver Attention Monitoring System (DAMS) sets out test procedures and test methods that were implemented in May 2023 [11]. The 5 types of distracting behavior are described, and associated warning conditions specified.

Emotion 3D summarizes the criteria for the recognition of specific behaviors mentioned in the regulation as follows [12]:

- **Eye-closing**

- Eyes completely closed for a duration of ≥ 2 s

- **Abnormal head pose**
 - Head deflection angle left or right $\geq 45^\circ$ for a duration of ≥ 3 s
 - Head up or down $\geq 30^\circ$ for a duration of ≥ 3 s
- **Answering hand-held phone**
 - Distance between any point of the hand-held phone and the face < 5 cm for a duration of ≥ 3 s
- **Yawning**
 - Mouth opening height-width ratio (the ratio of the minimum vertical height of the inner edge of the upper and lower lips to the horizontal width of corners of the mouth) > 0.6 for a duration of ≥ 3 s
- **Smoking**
 - The minimum distance between the hand-held cigarette and the lips shall not be greater than 2 cm, for a duration of ≥ 2 s

Japan

Safety regulations for automated operation devices

In March 2020, safety regulations for automated operation devices were formulated by the Ministry of Land, Infrastructure, Transport and Tourism (MLIT) within the Road Transport Vehicle Act (enacted in May 2019, took effect in April 2020) [13].

These safety regulations establish driver monitoring technology as equipment for vehicles that monitor the driver's condition and ensure whether they are ready to take over. In the following, main requirements are specified:

Performance requirements

1. Safe operation shall be continued until the driver takes over, and the vehicle shall be stopped safely if he or she does not take over. An alarm to alert the driver before taking over operation shall be made before leaving the standard operating environment.
2. The vehicle shall be equipped with driver monitoring to monitor the driver's condition.
3. Measures shall be taken to ensure cyber security to prevent unauthorized access.

Operation status recording equipment

- ON/OFF time of the automated operation device
- Time when the alarm was triggered to take over driving
- Time when the driver became unable to respond, etc. must be able to be recorded for 6 months (or 2,500 times).

India

The Ministry of Road Transport and Highways (MoRTH) has instructed the Automotive Industry Standard committee (AISC) with a final draft (Draft AIS-184/DF) in 2022 [14], to initiate

a new Automotive Industry Standard (AIS) with the topic of Driver Drowsiness and Attention Warning Systems (DDAW) that shall meet the European requirements (see Article 6 of the European Regulation 2019/144).

General requirements

- The DDAW system shall issue a warning to the driver when the drowsiness level is 7 and 8 or higher, according to the Karolinska Sleepiness Scale (KSS).
- Functionality of the DDAW system shall be without biometric information, facial recognition of the vehicle occupants. Any processing of personal data shall be carried out in accordance with data protection law.
- Manufacturers must carry out validation tests to ensure that DDAW systems are capable of monitoring driver fatigue accurately, robustly and scientifically.

While the draft calls for the indirect nature of the measurements and requirements to be considered, they should also be technology-neutral to encourage the development of newer technologies.

UNECE

UNECE has adopted a framework document in automated/autonomous vehicles by the World Forum of Harmonization of Vehicle Regulation (WP.29) at its 178th session [15]. The document presents the identification of key principles for ensuring consistent safety of Level 3 and higher automated/autonomous vehicles, with the aim of harmonizing vehicle regulations.

New UN Regulation draft

On February 1, 2024, the UNECE Special Working Party on Automated/Autonomous and Connected Vehicles (GRVA) adopted a new draft regulation laying down provisions for the approval of vehicles with Driver Control Assistance Systems (DCAS) and minimum safety requirements for vehicles equipped with advanced driver assistance systems (ADAS) [16]. Building on UN-Regulation No. 79 published in 2018, the new draft regulation includes an expansion of the technologies to be introduced in new models.

According to the regulation, DCAS, as a subgroup of Advanced Driver Assistance Systems, must guarantee and ensure through its design that the driver fulfills his driving task.

- The driver's vision (visual interaction between the driver and the road) must be monitored
- the driver's hands must remain on the steering wheel.
- An alarm must be triggered after 5 seconds if it detects that this is no longer the case.

The purpose of this is to prevent the driver from relying too much on such systems and overestimating them (Mode Awareness).

Framework document on automated/autonomous vehicles

During the 178th session of the UN WP. 29 (World Forum for Harmonization of Vehicle Regulations) in February 2022, a Framework Document on Automated/Autonomous Vehicles (FDVA) of levels 3 and higher was published [15]. This document was prepared at the interstate level by representatives from China, the European Union, Japan, and the United States of America. The aim of the framework is to harmonize worldwide requirements and guidelines for automated driving and to create space for the development of innovations. Updated tables show HMI (internal and external) and driver monitoring as functional requirements for automated/autonomous vehicles. New assessment and test methods demands for in-use monitoring.

1.2 Vehicle Assessment Programs and Consumer Tests

In addition to legal regulations, vehicle assessment programs e.g. from NCAP define how Interior Monitoring and Driver Monitoring Systems are tested and assessed. Such assessment programs define the priorities for the new car market in the next decade. Although they are not legally binding, high safety ratings are important arguments for customer trust, homologation and legal disputes. Consumer tests in Europe (Euro NCAP), America (IIHTS-HLDI), Asia (e.g. ASEAN NCAP and CNCAP) and Australia (ANCAP) have adapted their criteria in the course of technological turnaround.

EU-NCAP

The European New Car Assessment Program (Euro NCAP) is a European voluntary car safety performance assessment program.

Driver and in-cabin monitoring Systems are considered in the test protocols (see Euro NCAP Vision 2030) [17].

Active safety

Driver inattention and drowsiness

Since 2020, Euro NCAP has been evaluating the first generations of indirect attention assistants, with sensors in the vehicle interior analyzing steering and head movements, eye gaze, and eye opening. Starting in 2023, test protocols have been expanded to include more advanced requirements for driver monitoring systems.

- Driver impairment
Alcohol influence, sudden motion sickness, as well as stress and cognitive distraction significantly impair the driver in their driving task and performance. The addition of advanced image sensors and/or biometric sensors, along with the introduction

of advanced functionality requirements, minimizes the risk of accidents and is strongly prioritized in practical implementation. Systems for detecting stress and cognitive distraction are long-term focuses of future developments.

- Child Presence Detection (CPD)
In response to the challenging issue of heatstroke fatalities in children left in cars, Euro NCAP aims to incentivize systems capable of detecting the presence of living beings, especially children, and alerting responsible individuals. Euro NCAP already released its Child Presence Detection protocol in 2023, rewarding both indirect and direct technologies through score points [18]. However, starting in 2025, only direct detection systems for capturing life within the vehicle interior will be rewarded.

Passive safety

Occupant classification

Driver Monitoring Systems enable the detection and classification of vehicle occupants, thereby facilitating the enhanced execution of other safety functions. The following safety functions shall be promoted:

- Airbag deployment parameters
- Seat belt force limiters
- Enhanced utilization, guidance, and optimization of seat belts and headrests (posture monitoring)
- Advanced airbag deactivation
- Reliable information on vehicle occupancy for advanced eCall/dCall systems.

Autonomous driving

To promote progress of automated vehicles, Euro NCAP aims to make the following changes:

- Development of best practice recommendations for safety assessment
- Publication of transparent requirements for AVs, based on which voluntary safety assessments (VSA) can be conducted.

Euro NCAP focuses its efforts on the following activities that will require in-cabin monitoring:

- Alternative seating positions
- Restraint systems
- Emergency measures

ANCAP

In November 2022, the Australian New Car Assessment Program (ANCAP) published its key updates to the existing safety assessment system in the document ANCAPS Future View on Testing & Assessment: An expanding focus to 2030, which was developed in collaboration with its European partner, Euro NCAP, whose report Vision 2030: "A Safer Future for Mobility" provides both content guidance and the basis for the protocol update for the Australian and New Zealand new car market [19].

Both organizations, ANCAP and Euro NCAP, will update the protocols in a three-year cycle, starting in 2026. Since 2018, ANCAP has been based on the overall Euro NCAP evaluation system and focuses on four central evaluation pillars: safe driving, crash avoidance, crash protection and post-crash safety.

In the 2023 changes to reward direct monitoring systems, ANCAP is advocating the following points for the content of the star rating system of future protocols:

- Leverage driver and occupant monitoring technology to facilitate other safety functions, such as smarter restraint deployment
- Address new critical scenarios and emerging road safety priorities through advances in-cabin sensing, software and connectivity

ANCAP formulates its efforts more specifically in cabin monitoring to promote and optimize safe driving.

Monitoring the driver

The development of future protocols aims to evaluate more advanced systems that are more robust and efficient.

This will cover the detection of driver impairment, including alcohol, the use of occupant status information to adjust the vehicle's performance characteristics (active and passive safety).

Occupant classification

In future protocol revisions, biometric data will be integrated to optimize the performance of safety features through the utilization of occupant information. These features include, for instance, seat belt pretensioning, load limiters and airbag functions.

Detecting the presence of children

Starting in 2023, the ANCAP rating system will encourage both indirect and direct detection methods. However, a transition is planned for 2025, at which point only direct detection will be evaluated.

IIHS-HLDI

According to new US regulations, semi-automated vehicles must have safety features that support the driver's concentration on the road. To address these requirements for driver monitoring and attention reminders, the Insurance Institute for Highway Safety (IIHS) has developed a new test program to evaluate safety features in semi-automated systems [20]. The IIHS test protocols and safety ratings will be included in Consumer Reports' scoring of semi-automated driving systems as they become available. The ratings of safety features as good, acceptable, marginal or poor are intended to support technologies and their functional aspects that reduce intentional and unintentional misuse. Systems monitor the driver's gaze, head position or hand position to ensure an active and conscious driving state.

Ratings are based on the output and type of warnings that remind the driver to look back at the road or put their hands on the steering wheel in the event of inattention. Possible warning signals include bell tones, vibrations, pulsating brakes or pulling on the seatbelt, which are emitted via several channels and are intended to escalate as urgency increases and time passes. If the driver fails to respond, the system slows the vehicle to a stop and should initiate notification of the emergency services if necessary. This escalation level allows the driver to be locked out of the system for the remainder of the drive.

Summarized requirements for a good partial automation safeguard:

- Monitors both the driver's gaze and hand position
- Uses multiple types of rapidly escalating alerts to get driver's attention

Fail-safe procedure slows vehicle, notifies manufacturer and keeps automation off limits for remainder of drive.

China NCAP

A new C-NCAP Management Regulation has been published in 2024 from the China Automotive Technology and Research Center Co., Ltd. (CATARC) [21]. The regulation calls for an official implementation from July 1, 2024 onwards. Child presence detection (CPD) evaluation items are added to the Occupant Protection Section. The Active Safety Section has been modified by introduction of the assessment item of the driver monitoring system (DMS).

The Appendix L, chapter 6.4 specifies DMS test side requirements and scenarios.



Figure 3: Overview consumer test organizations

1.2 User Requirements

In 2023 we conducted a brain-writing focus group in order to understand user requirements and to collect user ideas for applications of in-cabin monitoring. The use of cameras in vehicles is stimulated by user's expectations of in-cabin monitoring. Users benefit and acceptance is key to successful applications. By involving users in our research and development processes, these expectations and wishes to future functions and applications can be made visible and taken into account.

Method

In total 11 test person of different ages (19 to 48 years), genders (4 male/7 female), educational backgrounds and professions took part and were guided through discussion by two experts for

in-cabin sensing systems. With the combination of brain-writing, headstand method, focus group and online collaboration tools, a methodologically established approach was chosen for generating and categorizing ideas.

The complexity of the technology as well as the heterogeneity of the participants required, prior to idea generation, the establishment of a common starting point and working basis. Thus, the technology and the potential applications of occupant state recognition via optical sensors were described with the metaphor of the butler, who observes situations and offers appropriate assistance when needed. This step ensures that all participants have a common understanding of the technology's capabilities as well as the study objective, validating the final quality of the results.



Figure 4: Metaphor for in-cabin monitoring systems as a butler

Results and conclusions

Clustered and prioritized by two technology experts, a total of 29 applications from the areas of active safety, passive safety, entertainment and comfort were obtained (see figure 5).

Apparently, users have a lot of ideas for applications and functions that can be implemented with cameras in the vehicle.

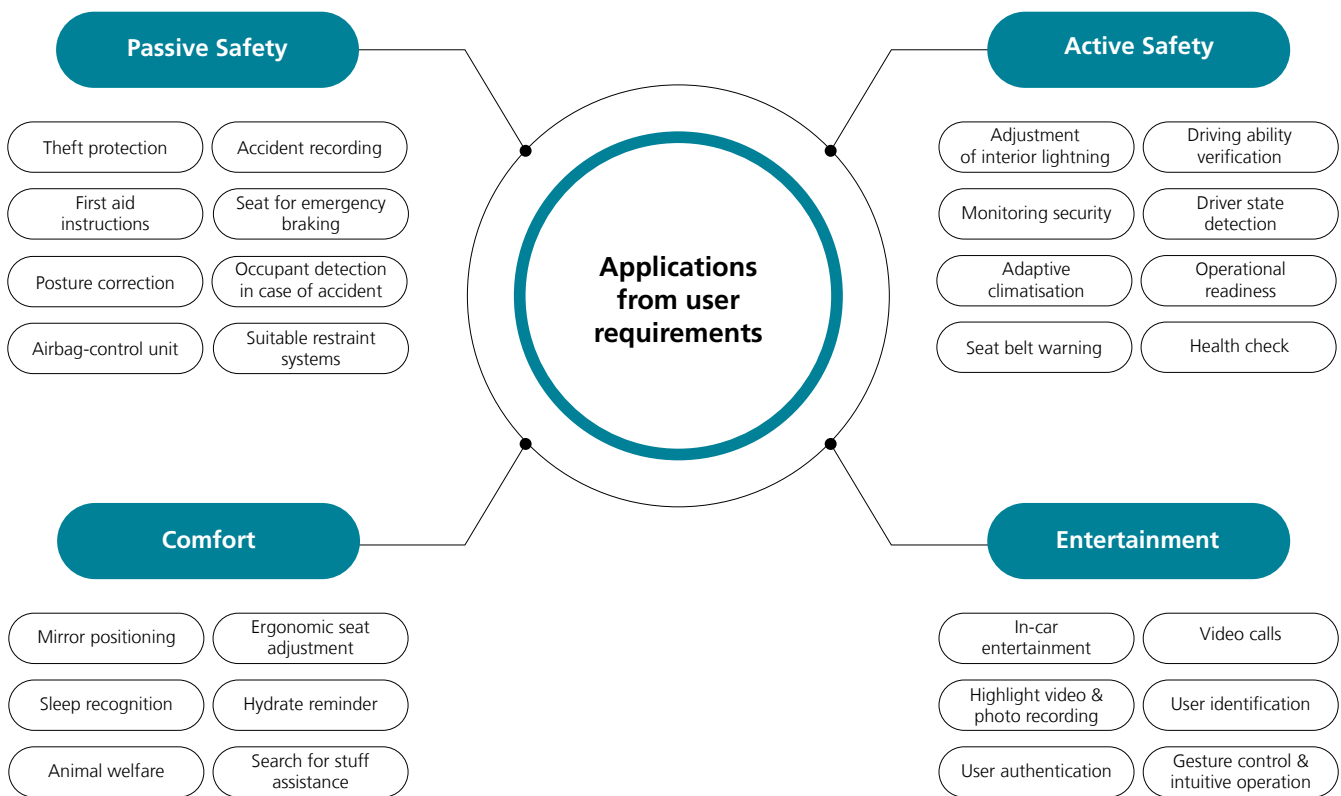


Figure 5: Overview of applications from user requirements study

1.3 Technology providers

Interior Monitoring solutions have arrived in the automotive industry. A market analysis reveals: Interior Monitoring is not just a trend but the future. The integration of Interior Monitoring is not merely a response to market demands; it's a proactive step towards shaping the future of vehicle safety and comfort. Guided by the societal and regulative demand for increased safety and comfort in vehicles, companies are aligning their portfolios with in-cabin sensing technologies. By doing so, they position themselves as technology leaders driving innovative and forward-looking developments in the automotive sector. Figure 7 provides an overview of in-cabin monitoring applications that are already on market. As an excerpt, the following companies already offer interior monitoring technologies.

The technology providers range from component developers, such as Infineon and Melexis, to algorithm providers such as SIMI, up to system integrators such as Bosch, Continental and Valeo.



Figure 6: Selection of companies working on DMS

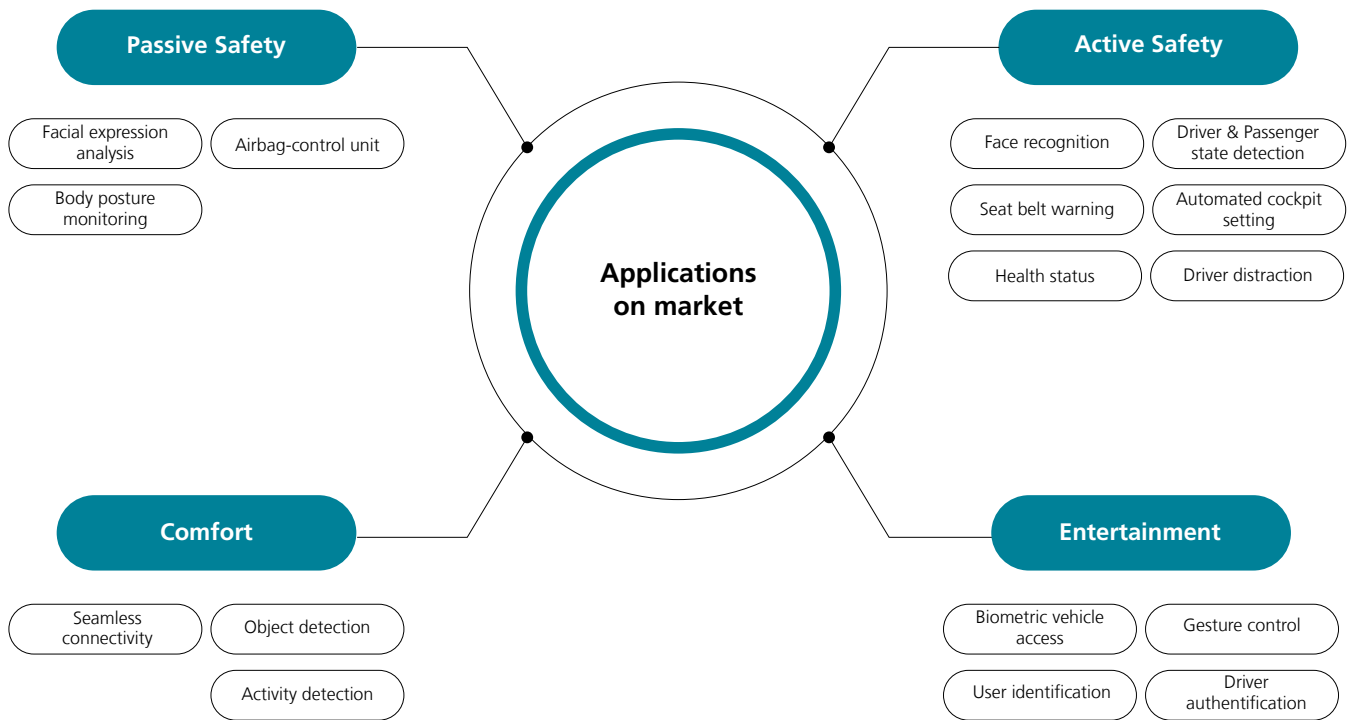


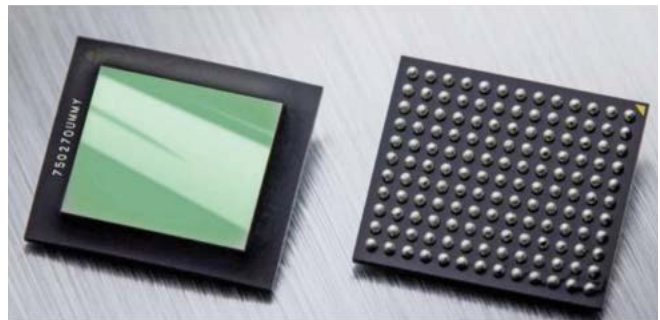
Figure 7: Overview applications on market

1.4 Realizing DMS and OMS Applications with a ToF sensor

Technology providers have been demonstrating high quality and advanced in-cabin monitoring applications with different sensors. In a joint project of Melexis and Emotion3D, both companies demonstrated a combined Driver Monitoring System (DMS) and Occupant Monitoring System (OMS), based on a single camera with ToF technology. The combined system is called Driver and Occupant Monitoring System (DOMS).

The project results were demonstrated at CES 2024 in Las Vegas.

The Melexis ToF sensor provides a full VGA resolution of 640 x 480 pixels for the depth image as well as for the NIR image. At full resolution this allows for a readout rate of up to 120 fps which is required for detecting fast movements.



- Full VGA 640 x 480 pixel resolution
- Full resolution readout up to 120 fps
- Simultaneous depth and NIR image caption
- Support up to ASIL B system integration
- Industry high quantum efficiency
- Flexible parameter set, incl. modulation frequency (up to 100 MHz) and exposure time
- 11 x 9.5 mm IBGA package

Figure 8: Melexis ToF sensor specifications



By combining depth and NIR images, EuroNCAP compliant DMS in combination with OMS applications are achieved with a single ToF sensor from Melexis. "

Kristof Lieben, Product Manager at Melexis

At CES 2024, visitors could experience the ambient light robust DMS applications, as required by EuroNCAP, based on the sensor inherent NIR image with VGA resolution. that are required by EuroNCAP, specifically:

- Distraction/Drowsiness/Attention classification based on
 - Eye-gaze
 - Eye-openness/blinking
 - Head position

Additionally DMS relevant feature detections were demonstrated for applications that use DMS:

- Calling on the hand-held phone
- Smoking
- Yawing
- Spoofing authentication
- Wearing glasses or a mask

Some of those activities are considered to be distracting and to be related to higher risks of accidents.

For the realization of OMS and passive safety applications, the Melexis ToF depth image was used in combination with the NIR image. Within a 110° x 85° field of view of the camera, following features could be detected, which are the foundation for restraint control systems and dynamic airbag control, as well as for comfort applications for the vehicle passengers.

- Sitting height
- Occupant classification
- Out of position detection
- High speed body pose tracking for safety applications.

The realization of the DOMS with only one VGA ToF sensor was achieved by combining the depth image and the NIR image of the Melexis ToF sensor.

- DMS functions are realized with the NIR image. The algorithms from Emotion3D required only little retraining to adapt to the inherent NIR image.
- OMS and passive safety functions are realized based on the 3D body pose which is extracted from a combination of the sensor's depth and NIR image.

The aim of the project was to realize a single camera system not only capable of doing EuroNCAP compliant DMS applications, but also OMS and passive safety.

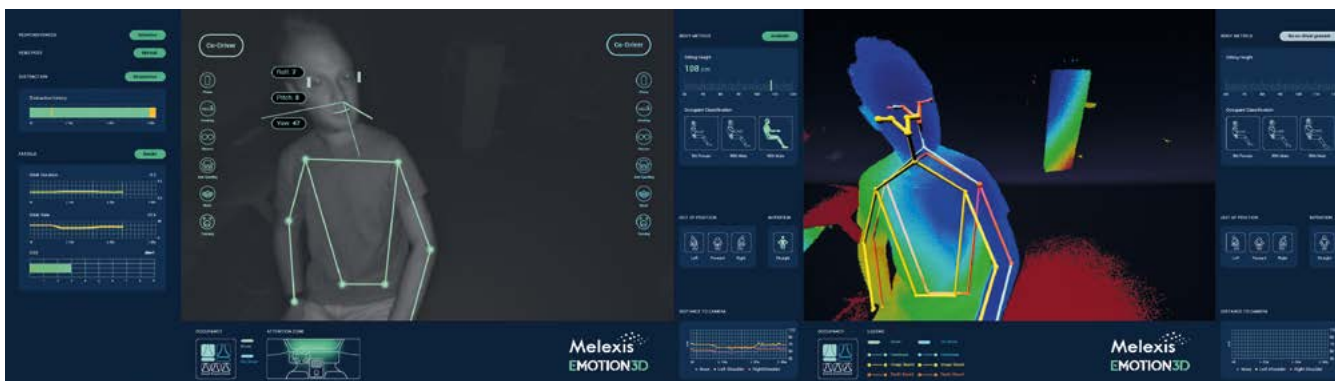


Figure 9: OMS detection features from combination of NIR and depth image



Figure 10: DMS detection features with Melexis ToF and its NIR image sensor

1.5 Consumer Electronics

Smart home and smart phone camera-based applications are not only stimulating but transformative for in-cabin monitoring. In-cabin monitoring enhances the user experience by elevating personalized comfort and entertainment for occupants. Inspiration for camera-based applications stems from the functionalities found in smart home and smart phone applications. Features like user authentication, face pay, memoji-based virtual chatting, and other smart home and smart phone applications can be migrated into camera equipped vehicles.

Through user status detection and identification, these applications can seamlessly enhance the in-car environment, providing a more connected and personalized experience in the vehicle interior. This integration opens possibilities for advanced customization, convenience, and entertainment options within the automotive space. The following camera-based applications are based on the possibilities of Smart Home and Smart Phone cameras.

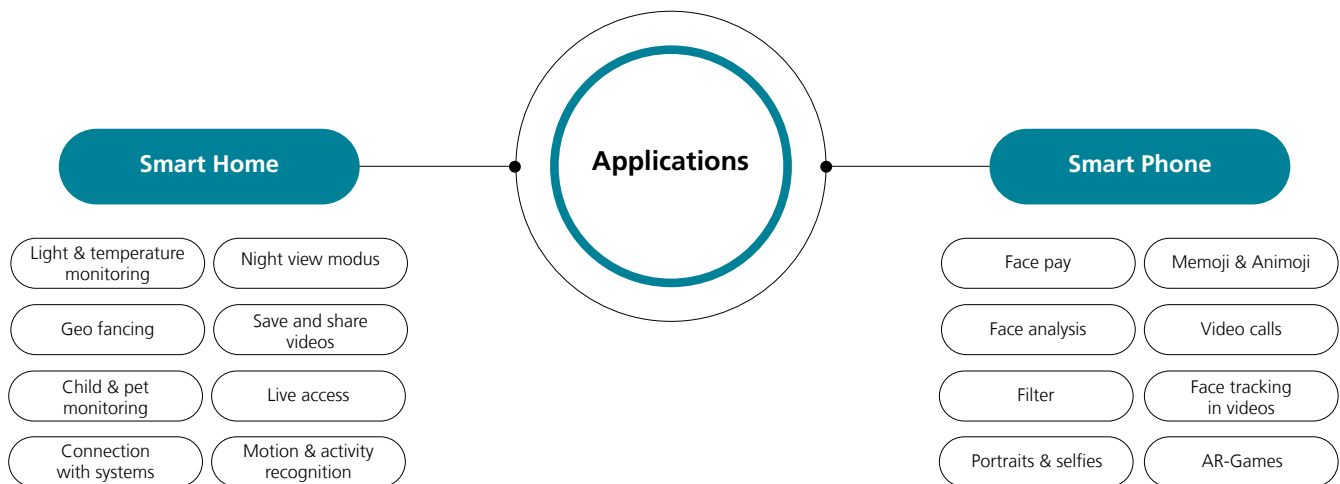


Figure 11: Overview applications from consumer electronics

1.6 Interview with Quirin Anker from Daimler Truck (FUSO) about DMS for Trucks

Quirin Anker has been involved in the development of driver monitoring systems for Daimler Truck and for features for the Japanese market in vehicles of the brand FUSO. He studied Automotive Engineering in Munich with a focus on powertrain, driving assistance and acoustics. His master thesis treated the evaluation of a driver's fatigue detection system. Since 2020 he is ADAS Engineer with a focus on driver monitoring systems. He has been involved in the development of local DMS solutions at FUSO for the Japanese market and in the development of global system solutions for Daimler Truck.



Figure 12: Quirin Anker from Daimler Truck (FUSO)

Japan is a leading country in introducing driver monitoring systems with camera technology first introducing such technology in the early years of 2000 and since then leading the market in passenger cars and commercial vehicles. What are the pull factors in Japan?

Several heavy accidents in Japan pulled the public and political attention to driver distraction and drowsiness. As a result, manufacturers pushed for a higher penetration rate in passenger cars and commercial vehicles and found a market willing to accept and pay for it. Due to the high acceptance rate in Japan, currently no legislative initiative is needed to push driver monitoring systems into the market. Contrarily, governments and authorities in other markets have been – rather recently – releasing regulatory frameworks to enforce a comprehensive introduction of driver monitoring systems, such as in the European Union or in the People's Republic of China. Camera-based driver monitoring systems are legally bound in the EU and People's Republic of China from 2024 onwards, whereas FUSO has been introducing its first generation in 2017 in Japan, called the Active Attention Assist. This feature is fusing driving behaviour-based parameters with an interior NIR camera.

What do your customers expect from driver monitoring systems (DMS) in trucks?

Our customers are mainly fleet owners and less private customers or end-users. Their buying decisions are driven by economic factors, such as fuel-efficiency, durability, longevity and of course safety. As a matter of fact, the human factors are still the dominant root cause of road accidents. Hence an effective, low-cost driver monitoring is what our clients ask for. Additional to this, the truck drivers care about having effective and highly accurate systems, which are not annoying and have a user-friendly interface. With the introduction of cameras in driver monitoring systems, accuracy and the number of applications could be improved significantly.



With the introduction of cameras in driver monitoring systems, the number of use cases and the accuracy and graduation of the warnings could be improved significantly.” Quirin Anker

What benefit provide DMS today?

DMS can be utilized in many ways, but for a commercial vehicle manufacturer the safety aspect is the most important. DMS can help the driver to stay more focused while driving, by warning and reminding the driver. Those use-cases can be related to monitoring over a longer period (driver drowsiness) or detecting imminent critical situations (micro-sleep or distraction). For example, DMS recommend taking a break when high drowsiness levels are detected or when a driver is distracted. With such use-cases the risk of safety critical driving situations and ultimately accidents can be reduced and mitigated.

Furthermore, DMS provide an indirect benefit for improving safe driving. Having DMS available and sharing warning occurrences with fleet owners could also help to create transparency and awareness of the working conditions as well as the driver's individual condition. In some cases this might have an impact on improving the working conditions and bring health benefits to the driver (e.g. better sleep quality due to better shift management).

Which camera technology is used to realize the desired functions?

The adequate camera technology depends on the target function. In most cases an active infrared illumination is required for stable image quality in all, but especially in low light conditions. A reliable 3D depth image can only be measured with Time of Flight or stereo cameras. This is needed in high precision, e.g. when measuring the drivers longitudinal head or body movements. Mono cameras may be good enough in lateral position however.

Camera position is an important criterion, especially if only one camera is used. An ideal camera position provides a non-occluded field of view without obstruction objects and at a suitable angle.

The field of view is also an important factor to consider, since algorithms, such as eye-tracking, require a resolution with sufficient number of pixels on facial features. Narrow field of view cameras are more cost effective for this use case. Other functions, such as video calls and cabin monitoring require a wide field of view.

Which functions do you see in the future and how would you cluster them?

Safety is the major focus on functions in commercial vehicles. This is reflected in regulations worldwide.

Effective systems also hold a high potential for improving Advanced Driver Assistance Systems and driving functions.

From my point of view this is of very high priority after fulfilling the legal regulations.

Video recording for liability cases, theft protection and training are additional functions.

Comfort functions however have a rather low-priority still for

commercial vehicle manufacturers. Gesture control features, video-calling etc. might be of interest, but only if it brings a clear benefit for drivers or fleet owners. E.g. reduced distraction or increased fitness to drive.

For fleet owners, generally every feature, that reduces cost or stand-still time, increases efficiency, reliability and safety is of high interest. I would cluster the functions as followings:

- Driver condition and health monitoring
- Driver distraction monitoring, with visual distraction and with distracting activities
- Interior monitoring and sensing
- Co-driver and passenger monitoring
- Infotainment and comfort functions

How do you see the future of in-cabin monitoring?

State of the art driver monitoring is a bridging technology for classic and assisted driving and will continue having its relevance in the market for the next decade to come, especially since it is tied to regulations (e.g. GSR DDAW & ADDW) and its revisions. The technology will contribute effectively to assisted vehicle control, e.g. by considering attention level, eye gaze and driver intentions in assisted and automated driving functions.

Developing active intervention concepts, countermeasure strategies and further use-cases based on the driver condition monitoring is the next step after fulfilling regulatory requirements and current market demands.

In commercial vehicles, factors such as reliability, efficiency, safety and practicability are among the most relevant factors for customers to choose a vehicle. Putting this into the context of driver monitoring, safety features and practicable features are the focus area for a CMV manufacturer.

Towards increased automated driving time, health monitoring, presence detection and comfort use-cases with in-cabin monitoring supported HMI might become more useful, since drivers would be less required for the actual vehicle control.

For example, the industry is already displaying first concepts of radar-based heartbeat monitoring, which could ultimately be used together with camera systems to monitor the health conditions of a driver. Such information could be used in the future for more precise health monitoring and initiating active steps in health critical situations (e.g. automatic safety stops and placing emergency calls. However, I currently do not see the radar-based technology to be ready soon.

Thank you, Quirin for this inspiring interview.

1.7 Marco Willems from CARIAD about Digital Twins

At the AutoSense InCabin Conference 2023 in Brussels on June 22nd Marco Willems, Head of Development Vehicle Occupant Monitoring at CARIAD presented a vision of creating digital twins of all cabin occupants.

Who needs the digital twin?

Marco Willems pointed out that in the generation of Software Defined Cars and Human Centred Cars, this digital twin representation is a core component to enable the next evolution of comfort and safety features for today's and tomorrow's consumers.

The best representation of a digital twin will enable brands and service providers to differentiate from competitors and enable the best and most personalized user experience on the market.

Hence the ultimate goal is a rich digital representation of all cabin occupants, with traits and states, that other agents in the car can use as a source of data.

What is the digital twin?

Marco Willems envisions a holistic representation of the occupants in a digital twin. He states that CARIAD will use a fusion of sensors for the in-cabin monitoring. As it is done today for exterior monitoring. He states that relying on only one sensor, leaves too many blind spots. Sensor fusion is the key. It is also most cost efficient.

According to Marco Willems, the digital twin appears to be an open-ended feature list of the occupant's traits and states and other data. He named some examples:

- Traits:
 - Weight
 - Height
- Constructs of cognitive states:
 - Mood
 - Workload
 - Cognitive focus
 - Drowsiness
- Physical state:
 - Body Pose
- Physiological data:
 - Heart rate

Who wants the digital twin?

According to Marco Willems, occupant monitoring is not a function for itself, but an enabler for many other functions and services in the car.

Today a car is designed in a "one size fits all" approach. In the future of digital cars, this will change.



**CARIAD Mission:
We try to create a holistic interior scene understanding, and to give this information to all features and assistants in the car."**

Marco Willems



From cars optimized for a standard person, towards highly individual user experience in the car."

Marco Willems



You have to know about the humans in the car."

Marco Willems

According to Marco Willems we see three stakeholders with requirements towards highly personalized and universal digital twins derived from occupant monitoring.

Legislation will not accept discrimination against ethnic groups by treating them differently. The representation has to work for everybody.

Society will expect safely driven vehicles and same value for money for everyone, fostering diversity and personalization. Some customers will not accept being monitored all day. And others want to see the technology which takes care for them very present.

The contradicting customer expectations appear to be related to camera integration, but may also play a role how obvious assistants use information from the digital twin.

Who uses the digital twin?

According to Marco Willems, occupant monitoring is not a function for itself, but an enabler for many other functions and services in the car.

Occupant monitoring is an enabler for many functions that customers buy:

He points out five categories of functions that can be enhanced with digital twins:

1. Assisted and automated driving
 - a. Is the driver attended to the road?
 - b. Is the driver ready to take over control?
 - c. Do the occupants behave in the way they have to behave?
2. Active safety systems
 - a. Attention
 - b. Occupant pose

- c. Identification
- d. Workload
- e. Cognitive state
- f. Side activities

3. Passive safety systems, especially in open, non-traditional interior settings
 - a. Position
 - b. Identification
 - c. Child detection
 - d. Object detection
4. UI/UX and entertainment, incl. new digital business models
 - a. Gaming
 - b. Gesture control
 - c. Identification
 - d. Mood
5. Health and well-being
 - a. Tele-medicine in the car
 - b. Health data collection

What are the challenges of an OEM?

While an OEM can buy occupant monitoring or entire systems that deliver a digital twin, it still has to solve challenges with regards to integration, good enough quality, legal compliance and business models. Marco Willems presented the pyramid of challenges where CARAID is aiming for the best sweet spot.

The four corners of the pyramid are composed of:

1. Lean Design
2. Best Function
3. Security
4. Costs

In Summary, Marco Willems emphasizes the need of occupant monitoring system for future cars.



Advanced Occupant Monitoring enables future car concepts, innovative features and unlocks a huge data source for digital business models in the future."

Marco Willems

Following chapter 1, a comprehensive overview of the identified applications sourced from user requirements (chapter 1.2), technology providers (chapter 1.3 and 1.4),

and consumer electronics (chapter 1.5) is provided. The applications are categorized by their source and respective vehicle categories.

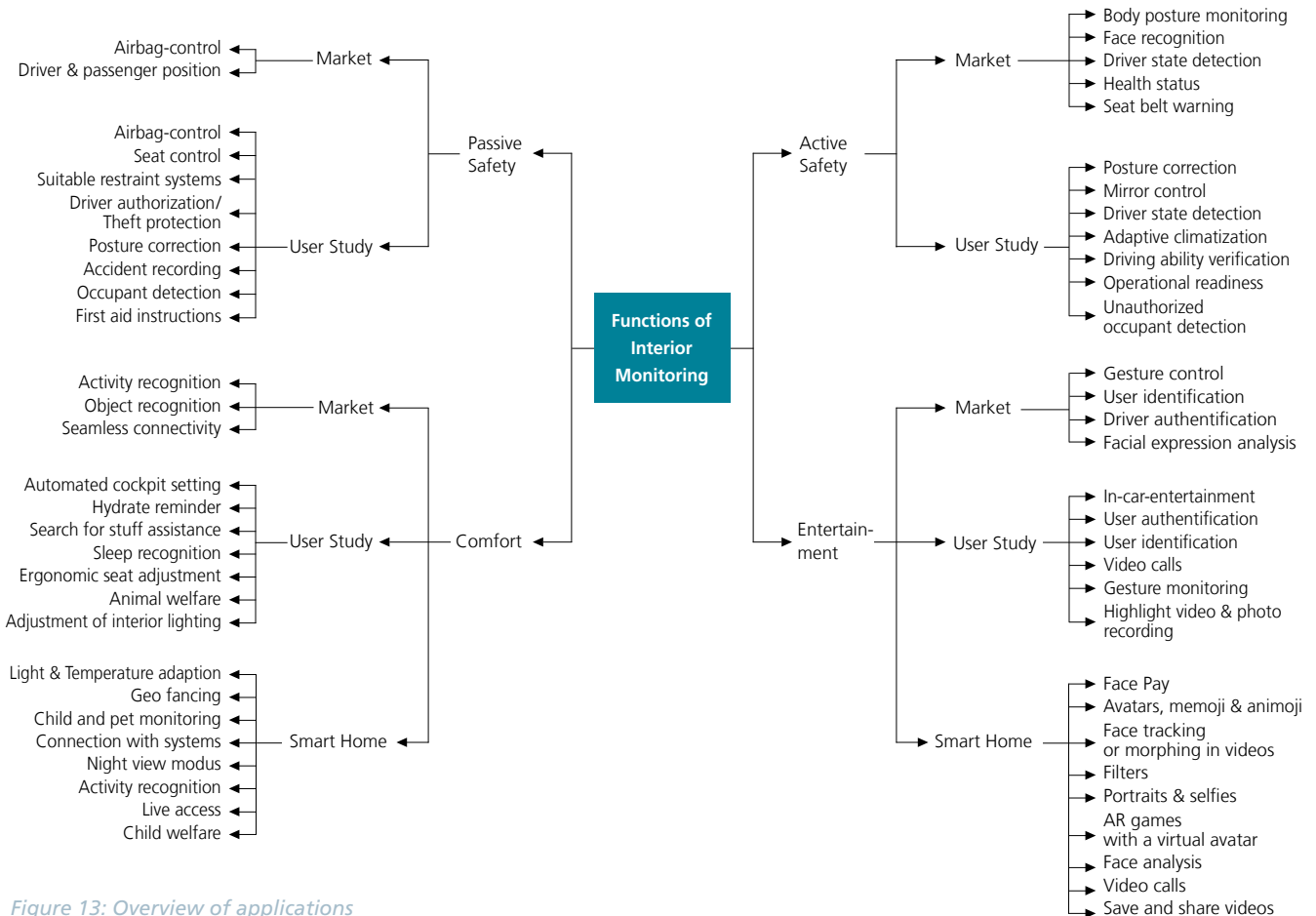


Figure 13: Overview of applications

2. Applications for In-Cabin Monitoring

Chapter 1 collected applications for interior monitoring from different sources, such as

- Legislation
- Consumer test programs
- Technology providers
- User requirements
- Other domains such as smart home and smart phone

Chapter 2 summarizes and specifies the applications.

Further we categorize the applications into the areas of:

- Active Safety
- Passive Safety
- Entertainment
- Comfort
- AI personal assistants

In the subsequent chapters, the detection functions are individually introduced and defined based on our experience and knowledge.

2.1 Active Safety

Under the chapter **Active Safety**, all applications of assisted and automated driving are compiled with the aim of improving safety and preventing accidents.

Active safety refers to assisted and automated driving that recognizes and warns about potential dangers for occupants. It may also steer or brake in order to reduce accident risks. With the aim of improving safety and preventing accidents, corrective measures are taken, such as controlling the vehicle in the event of inattentive drivers.

Driving ability verification | Verification of the driver's condition to ensure fitness to drive. To intervene in cases of inadequate states such as alcohol-impaired driving, fatigue or illness. Also ensuring fitness to drive with respect to constant traits such as legal age, driver's license, etc.

Source: Market

Unauthorized occupant detection | Detection of attempted break-ins or unauthorized vehicle access. Logging and transmission of the recordings to the owner and relevant authorities, e.g. police, insurance company.

Source: Market

Driver state detection | The interior cameras record and monitor the current state and activity of the driver while driving. In this way, risky situations such as tiredness, distraction or emotional states can be detected.

Source: Market

Adaptive climatization | Adjustment to personalized and preferred climate settings. The aim is to optimize driving comfort for each passenger, minimize energy consumption and generate a pleasant environment in the vehicle.

Source: User Workshop

Mirror control | Head- and eye-position tracking allows an automated and situation-specific adjustments of rearview mirrors. Digital mirrors can respond to the viewing angle and increase the field of view according to head movements.

Source: User Study

Operational readiness | Detection and ensuring that the vehicle and the interior configuration is ready for use and able to perform its intended tasks safely and effectively.

Source: Market

Seat belt warning | Recognizing and warning of improper or non-use of the seat belt. Warning system for the correct use of belts.

Source: Market

Health status | Recognize signs of illness leading to immediate loss of fitness to drive such as a stroke or in the event of an accident, scan the injuries and issue instructions for action.

Source: Market/User Workshop

Posture correction | Automatically adjusting the driver's seat and controls to ensure optimal ergonomic positioning, reducing fatigue and the risk of musculoskeletal issues.

Source: User Study

2.2 Passive Safety

Passive Safety refers to functions that protect occupants from injuries during accidents. Passive safety typically aims for protecting the human body.

Driver & passenger position | Recognition, counting and location of all occupants positions in the vehicle. This can help passive safety systems to protect passengers in all seats.

Source: User study, Market

Airbag control | Detection of seat occupation, estimation of body pose and safe seating position, location of body parts and objects and estimation of occupants' weight.

Source: Market | User Study

Seat control | A specially designed car seat that adjusts the occupant's position during sudden braking and before crashes to reduce the risk of injury.

Source: User Study

Accident recording | Utilizing onboard cameras and sensors to record data and footage before, during and after a collision, providing crucial information for insurance and legal purposes.

Source: User Study

Occupant detection | in case of an accident the system detects the presence and position of passengers in the vehicle to inform emergency response in the event of an accident.

Source: User Study

First aid instructions | Providing automated, voice-guided first aid instructions through the vehicle's infotainment system in the event of an accident or medical emergency.

Source: User Study

Driver authorization/theft protection | The permission to drive the vehicle and to be responsible for moving it, can be controlled by driver authentication. Theft and misuse can be hindered.

Source: User Study

2.3 Entertainment

The Entertainment chapter brings together functions for entertaining occupants and improving the human machine interface (HMI). Entertainment functions typically aim for increasing user experience for the car occupants.

Facial expression analyses | Interpretation of human facial expressions, often for the purpose of understanding emotions or reactions. Human Machine Interaction, entertainment and content can be adapted to facial expressions.

Source: Market

Gesture control | A human machine interface allowing humans to operate certain vehicle functions such as adjusting the stereo or climate control through specific hand movements.

Source: Market/User Study

User identification | Recognizing a user and distinguishing the user from other users in the car, e.g. distinguishing between driver and co-driver, to adjust vehicle HMI, settings and permissions.

Source: Market/User Study

Driver authentication | Ensuring that the person operating the vehicle is authorized to do so, or authorize payments.

Source: Market

Video calls | The ability for real-time video communication between passengers or drivers in the vehicle and external contacts. Typically, the camera capturing users inside the vehicle is combined with a display to see the external person.

Source: User Study/Smartphone

Portraits and selfies | The ability of the vehicle's infotainment system to take and save high-quality photos or self-portraits of passengers, often enhanced by lighting or background adjustments.

Source: Smartphone

Filters | Visual or acoustic effects applied in real time to video calls, photos or recordings within the vehicle system, often for entertainment or privacy purposes.

Source: Smartphone

Highlight video and photo recording | An automated function that captures and records important moments, memories or scenery while driving, using the vehicle's on-board cameras.

Source: User Study

Avatars, Memoji & Animoji | Customizable, animated characters or emojis in a vehicle's infotainment system that can

mimic the facial expressions of the driver or passenger and are often used for messaging or entertainment.

Source: Smartphone

Face tracking or morphing in videos | A feature that recognizes and tracks faces in real time during video conversations or recordings, ensuring consistent framing and focus. Also morphing the face towards the camera is helpful in the car, to create the perception that the passenger is looking directly into the camera, even though he or she may look on the road.

Source: Smartphone

AR games with a virtual avatar | Augmented reality games integrated into vehicle systems that provide interactive, immersive experiences and can incorporate the in-cabin situation and passenger movements into an avatar in the virtual world.

Source: Smartphone

2.4. Comfort

Comfort functions enhance the well-being of occupants or ease their achievement of goals. Some comfort functions can also improve users' health and fitness.

Automated cockpit setting | Automatic and personalized adjustment of the driving seat position, steering wheel position and air conditioning to the driver's preferences when the occupant gets in. Radio stations, music, navigation and connectivity options can also be customized.

Source: Market

Adjustment of interior lighting | Changing the interior lighting to support passenger activities such as driving, searching, reading or sleeping or adapt the interior lighting to moods or circadian rhythm.

Source: Market

Health check | Monitor health over a longer period. Collect a baseline of health status and detect abnormalities. Perform early disease screenings.

Source: Market/User Workshop

Activity recognition and prediction | Many comfort functions can be realized or supported by activity recognition and prediction. A system that knows the current or future activities of car occupants can adjust automated driving performance, morph interior settings, tailor speech dialogues and proactively support passengers. It can contribute noticeably to the user experience and personalization.

Source: Market/Smart Home

Object recognition | Recognizing objects in the vehicle interior can support many comfort functions. A system that knows which objects are in the car, can help searching for it, provide an inventory of (forgotten) objects and adapt for the context when such systems are used.

Source: Market

Hydrate reminder | A system within the vehicle that reminds occupants to drink water at regular intervals, promoting hydration and overall health during long drives.

Source: User Study

Sleep recognition | Detecting sleep of passengers can improve comfort by adapting driving dynamics, noise cancellation, temperature control, and light control. Wakeup scenarios can be trigger with personalized snooze time.

Source: User Study

Ergonomic seat adjustment | Automated adjustment of the vehicle's seats according to the occupant's body shape and size for optimal comfort and posture, reducing fatigue and pressure points.

Source: User Study

Animal welfare | Reminding owners of pets to care for them, e.g. not forgetting them alone, providing water and assuring safe and comfortable transportation.

Source: User Study

Light & temperature adaptation | A system that automatically adjusts the vehicle's interior lighting and climate based on occupant preferences for optimal comfort and visibility.

Source: Smart Home

Child presence detection/Child welfare | Reminding parents to care for children, e.g. not forgetting them alone, assuring safe and comfortable transportation, improving interaction between parents in the front and the child in the back.

Source: Smart Home/Market User Study

2.5 AI Personal Assistants

In-cabin monitoring with camera can contribute to the personalization and contextual adaptation of AI assistants in vehicles. This field is relatively new and holds a huge potential for many applications.

Adaptation of speech dialogues to visible information about past and current context, user state and about user activities will make speech dialogues much more intuitive, relevant and personalized. Information becomes more trustful if it is tailored to the current situation. Context awareness of AI assistants can also support explainable AI methods and dialogues.

Applications for AI assistants with OMS features are:

Pro-active intervention | Pro-active intervention at the right moment is possible when considering the passenger activities. Context sensitive pro-active support or approaching users' in-between activities can increase acceptance and flow.

Inventories | Inventories of objects seen in the car will allow to ask the car for support to find stuff (e.g. I saw an umbrella in the car yesterday, maybe it is still there?)

3. Roadmap for In-Cabin Monitoring Applications

The stakeholders pushing the introducing of in-cabin monitoring functions are in accordance to chapter 1:

- Legislation
- Test protocols
- User expectations
- Technology providers
- Users

With respect to the timing of introduction, the legislation provides ambitious deadlines for safety functions that require in-cabin monitoring. In accordance to legislation, also test protocols call for effective safety functions.

In the slipstream of safety functions, users and technology providers propose an armada of entertainment and comfort functions which become reachable with cameras in the interior. Also multimodal interaction with vehicle functions, e.g. by gestures will spread more and more. Large Visual Models that interpret visible scenes will significantly improve individualized interaction. Health functions are on the horizon with an enormous promise for user benefit and monetarization possibilities.

We foresee three waves of in-cabin monitoring integration:

1. Legislation drives the safety functions and is currently dominating development resources. Legislation does not require a special technology for the requested functions and 2D and 3D sensors have been used by technology providers to demonstrate the requested functions. Test protocols benchmark functions and will disclose which technology enables the best function performance. Driver fatigue and distraction monitoring has mostly be demonstrated with 2D cameras, but also ToF sensors have been demonstrated.

Legislation may benefit certain technologies, e.g. detection of children left alone in the car requires sensors that cover all parts of a car, including footwell on the backseat. Radar sensors are proposed for this use case, because seats, clothes and other objects in the car do not block the radar waves and hence do not occlude hidden places in a car. Smart airbags that adapt to the occupants' distance and body pose require depth information with a certain functional safety level. This is best provided by ToF cameras.

2. Followed by the safety functions, users request a variety of entertainment and comfort functions, that they know

and use in smart phones and smart homes. In those domains, high resolution RGB cameras are dominating and are used for photos, videos and video calls. ToF cameras are used in high-end smartphones for precise depth measurements, which is used for face authentication, depth effects in photos and AR visualizations. All those functions are meaningful in the car interior as well and user will ask for it. If it is not well integrated, they will use smartphones and after marked devices – with a burden of uncontrolled distraction and consequently safety impact. In parallel to the before mentioned functions, an intuitive and distraction free interaction with the vehicle is an important driver for in-cabin monitoring. Gesture recognition, activity recognition, body pose detection and occupancy provide a high benefit for context sensitive human machine interfaces. AI Assistants that are able to reflect complex user inputs, not only by text, but also by body language and facial expressions are likely to become more context sensitive, distraction free and more personalized by integrating input from in-cabin monitoring systems. The technology for this is not yet defined, but most likely a fusion of different optical sensors will provide the most complete digital representation of the human. Hence 2D, 3D and thermal optical sensors can provide this.

3. Health functions based on optical sensors are still under research and may be the latest wave for introduction in the car. A strong push from legislation, to detect intoxication and sudden incapacity is pushing such functions. Most research in this field is done with thermal cameras. Also precise 3D body pose and body movement measurements are useful for health monitoring.

4. Cameras for In-Cabin Monitoring

Many applications for interior monitoring are realized with some kind of optical sensor. Different optical sensors have been used, investigated, or proposed. Each optical sensor has specific characteristics and is more or less capable for one or the other application.

The challenge for automotive OEM is to find the best mixture of relevant applications and suitable sensors to achieve a high or good enough quality of the application.

4.1 What kind of camera sensors are suitable?

Cameras are small and cheap, non-invasive sensors. Different kinds of camera allow for different information extraction and a multi-view camera system can perform as a unit with broad perception possibilities.

Cameras provide information about

- depth
- surface reflectance and
- temperature.

Depth cameras like ToF, stereo systems, or multi-camera systems with triangulation provide the 3D shape or geometry of the scene.

2D methods with or without active illumination like RGB or NIR enable deductions about the scene's surface reflectance.

Thermal imaging sensors like FIR measure surface temperatures.

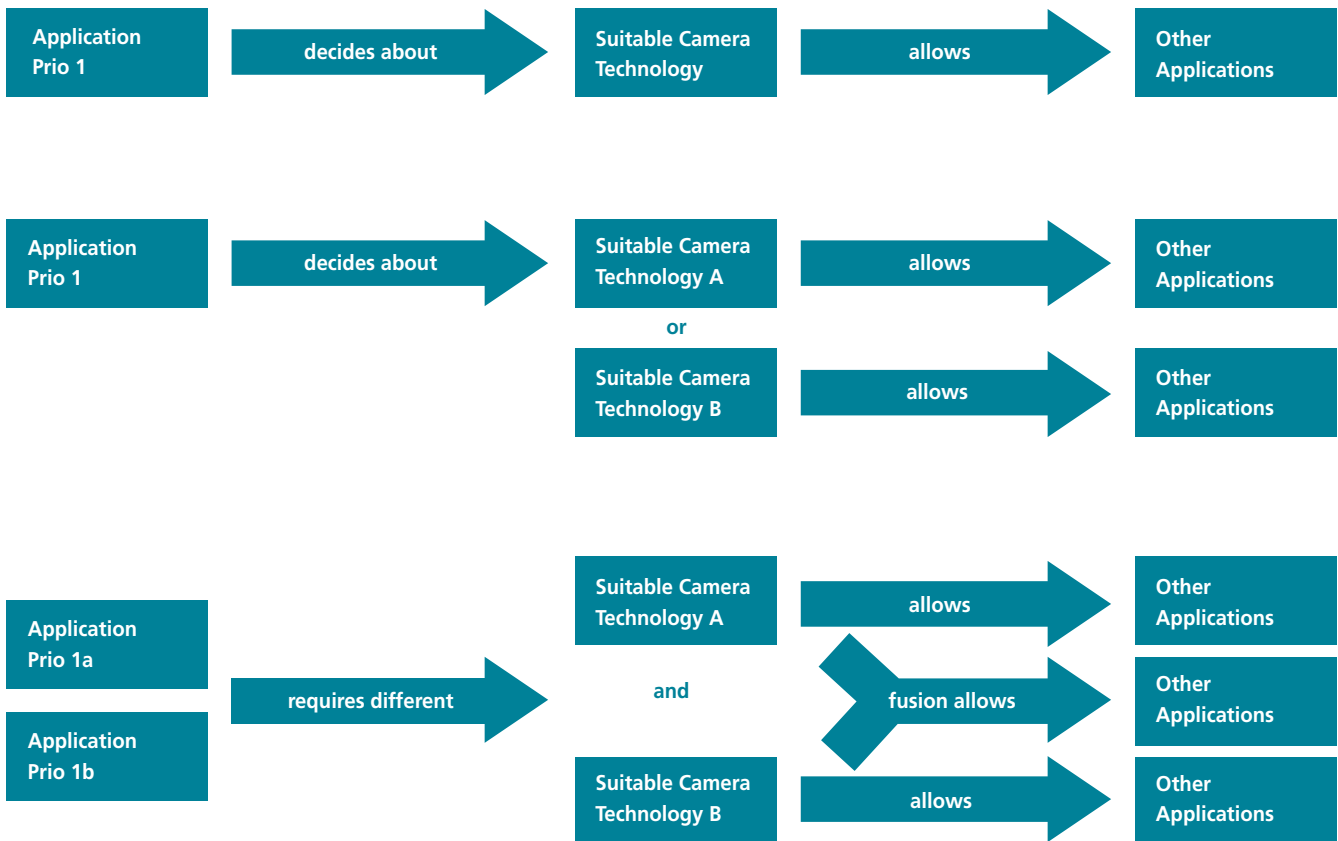


Figure 14: Start with the prio 1 application to decide on the camera technology which will enable further applications.

Each camera technology has its own advantages and disadvantages. There is no technology that is universally suitable for all applications. It is therefore important to select the most suitable camera technology for the prio 1 application. This camera technology will then enable further applications. Many times, more than one camera technology will be good enough to realize the prio 1 application. In this case, the other possible applications may drive the decision for the camera technology. And some times, different prio 1 applications will require different camera technologies. The fusion of different camera technologies opens the stage for the largest variety of applications. The combination of several camera technologies and even other sensors may significantly improve functionality and range of possible applications. By combining different technologies, gaps in the capabilities of one technology can be compensated by another technology. On the other hand, the usage of a single camera restricts the number of realizable functions.

Key differences of the sensors

- **Applications:** Each type of sensor is suited to different applications based on their ability to capture different types of information (color, temperature, light intensity, etc.).
- **Spectral sensitivity:** RGB captures visible light, NIR captures near-infrared, FIR captures thermal radiation, and monochrome captures light intensity without color.
- **Image content:** RGB produces color images, FIR produces thermal images, NIR can produce images based on infrared illumination (usually monochromatic), and monochrome produces grayscale images.

Radar: A radar system uses radio waves to detect the distance, speed, and direction of objects in the environment. Radar can work in a wider range of light conditions and passes objects in the car, such as seats and other passengers. It does not suffer from occlusions as most other cameras. It is not effective at detecting fine details such as facial expressions. Radar image quality also suffers from vibrations in moving cars. It is better suited for use in parking cars.

Time of Flight (ToF) Camera: Time-of-flight technology is based on measuring the time it takes for a light signal to travel from a light source to an object and from there back to the sensor surface.

A light pulse, usually in the infrared range, is generated by a light source, often a laser. This pulse is emitted in the direction of the object to be measured. The emitted light hits the target object and is reflected by its surface. The time it takes for the light to complete this round trip depends on the distance between the sensor and the object. A sensor catches the reflected light. The sensor measures the time it takes for the light to return. This is known as the time of flight.

By measuring the time, it takes the light to travel back and forth, the sensor can accurately calculate the distance between itself and the target object. The advantages of time-of-flight

sensors lie in their accuracy and fast detection of distances. They are well suited for applications such as distance measurement, gesture recognition and three-dimensional environment mapping.

Far-infrared (FIR) Camera: FIR (Far-Infrared) cameras are designed to detect and analyse infrared radiation, which is a type of electromagnetic radiation, invisible to the human eye but can be sensed as heat by humans. FIR cameras detect the infrared radiation emitted by objects and body parts. Every object emits some form of infrared radiation, and the amount varies with temperature. The human body, if warmer than the surrounding environment inside a car, emits infrared radiation detectable by FIR cameras. Occupant detection and changes in body temperature due to emotions or illness are applications for FIR cameras.

Since the system works on infrared, it can operate in low-light conditions or even complete darkness, making it effective at night or in tunnels.

Near Infrared (NIR) Camera: The NIR sensor emits invisible light pulses in the near infrared spectrum. These light pulses are invisible to the human eye. The material properties influence the reflection and absorption of the NIR light. The NIR sensor detects the reflected light that comes back from the surfaces of the objects. The reflected intensity and the spectrum of the light provide information about the material properties of the surfaces.

NIR sensors are often used for interior monitoring. They can monitor drivers' facial expressions and detect the direction of gaze. In addition, NIR sensors can be used in combination with other sensors for occupant monitoring and access control. The advantages of NIR sensors lie in their ability to provide information about surface conditions in the invisible NIR range. These sensors are able to detect subtle nuances in the reflection of light and are therefore well suited for applications that require precise and non-contact sensing in light and dark conditions.

RGB: RGB stands for red, green, and blue, which are the primary colors of light. An RGB sensor captures light in these three color channels. In an RGB camera, each pixel on the sensor is covered by a red, green, or blue filter. This setup allows the sensor to capture the intensity of each colour at every pixel, which can then be combined to produce a full-colour image. Downsides of RGB cameras are that they provide images only with visible light spectrum and hence cannot be combined with invisible infrared illumination. They do not work at night.

Table 2 maps the sensor functionality and operating principle with typical detectable functions with this sensor and typical applications for in-cabin monitoring.

Sensor	Functionality	Operating principle	Detectable features	Unique selling point/ Unique application area/ optimal for applications, such as
Radar	uses microwave or radio waves	measures time delay and Doppler shift	<ul style="list-style-type: none"> ■ Object detection ■ Object localization ■ Distance measurement ■ Movements 	<ul style="list-style-type: none"> ■ measures breath and heart rate ■ Child presence detection
Time of Flight	uses light pulses	measures the time it takes for light to travel from the source to the target and back	<ul style="list-style-type: none"> ■ Object detection ■ Object localization ■ 2D Body pose detection ■ 3D Body pose detection ■ 3D scene reconstruction ■ Movements e.g. steering wheel rotation ■ Distance measurement ■ Face recognition* ■ Eyetracking* 	<ul style="list-style-type: none"> ■ 3D sensing for passive safety ■ 3D Position of hands, head, hips ■ Smart airbag
FIR	detects thermal radiation	measures the intensity of light in the thermal wave length	<ul style="list-style-type: none"> ■ Body temperature 	<ul style="list-style-type: none"> ■ Occupant detection ■ Body temperature changes
NIR	detects near-infrared light, especially reflections from own active illumination	measures the intensity of light in the 850–940 nm wave length, typically combined with active illumination for night use	<ul style="list-style-type: none"> ■ Object detection ■ 2D Body pose detection ■ Movements e.g. steering wheel rotation ■ Face recognition ■ Eyetracking ■ Visual focus of attention 	<ul style="list-style-type: none"> ■ Eyetracking ■ Visual focus of attention/distracted

*at DMS position

Table 2: Comparison of sensors

4.2 Interview with Elena Zhelondz from A2MAC1 about In-Cabin Sensors in Series Production Vehicles



Figure 15: Elena Zhelondz, A2MAC1

Elena has a solid background in Electrical and Information Technology acquired through dual studies at the prestigious Technical University of Munich and CentraleSupélec, and she possesses a deep enthusiasm for technological advancements. Elena started her career at A2MAC1 in 2020, where she continues to apply her expertise as a Senior Cost Engineer in the dynamic field of electronics and autonomous driving.

With offices and teardown centers all over the world, A2MAC1 Automotive Benchmarking is a global reverse engineering firm directly serving every major automotive OEM, as well as 120+ members of the global automotive supply chain. A2MAC1 deploys its database in pursuit of identifying best practices and cost and design optimizations. Relying on this extensive database, A2MAC1 engineers are able to calculate costs of all electronic and mechanic components of a vehicle, as well as provide valuable technology insights and benchmarks.

Thank you Elena for the insight into the A2MAC1 database. A2MAC1 has torn down – not only for us – all modern car models worldwide to analyse the in-cabin monitoring cameras in modern vehicles.

What kinds of camera sensors do we find in today's cars?

In our database, the first cars with interior cameras started appearing in 2018. Over the years and with more safety norms being introduced, the share of vehicles with interior cameras has increased.

In general, we see Chinese OEMs setting the trends for in-cabin monitoring. Driver monitoring cameras are seen more often in Chinese vehicles, and only in cars made by Chinese OEMs do we see two interior cameras – one for driver monitoring and one for interior surveillance. In American and European cars, the introduction of occupant monitoring sensors is pioneered by upper-class vehicles.

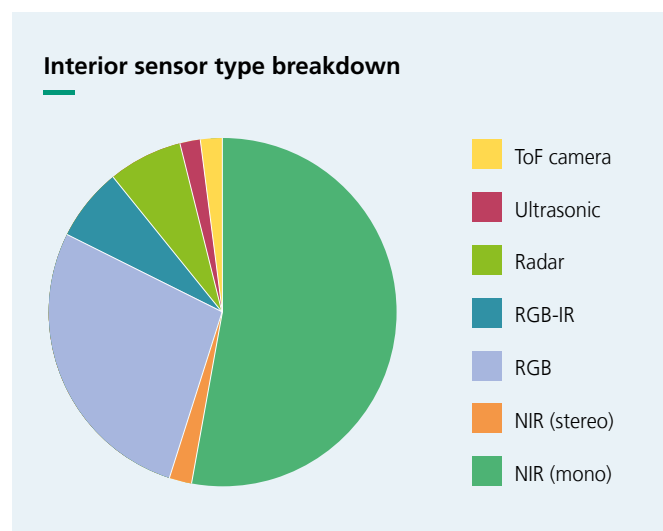


Figure 16: Sensors used for in-cabin monitoring



2D sensors are prevalent among in-cabin monitoring modules. While a 3D video feed provides more information, the additional hardware requirements imply a higher cost.”

Elena Zhelondz A2MAC1

As we can see on Figure 16, more than half of the in-car sensors are near-infrared (NIR) illuminated monochrome cameras. This means that several infrared LEDs are included in or next to the camera to illuminate the driver's face. Facial expressions and eye movements thus become visible even at night or with the driver wearing sunglasses. The use case for these is to issue alerts in case of driver fatigue or distraction.

The second largest share are regular RGB cameras that are usually mounted in the middle of the vehicle and oversee the whole cabin. These cameras have no illumination of their own and therefore cannot function well in the dark. The main use case is currently for infotainment, like taking photos or videos of the vehicle's passengers.

RGB-IR cameras take a step towards safety-relevant cabin monitoring by introducing a NIR component to an RGB camera while keeping color fidelity. On the hardware side, this means adding IR illumination to the camera. Similarly, to RGB sensors, RGB-IR cameras usually oversee the whole cabin.

Recently, non-camera-based sensors for passenger monitoring have been introduced, with either radar or ultrasonic signals as the enabling technology. These sensors lose the infotainment component of recording video material but can still reliably detect passengers and are very cost-efficient.

What are the main locations?

Driver monitoring cameras (DMCs) are used to notice if the driver is tired or distracted. They use active infrared illumination to ensure driver visibility in every setting and are focused on the driver. As seen on Figure 17, the most popular location is on the steering column or around the cluster (about 60 % of DMCs), followed by the A-pillar (25 %) and the rear-view mirror (7 %).



Figure 17: Most popular locations of DMCs

Interior surveillance sensors are usually positioned in the centre of the vehicle and as high as possible to enable optimal visibility of all occupants. RGB cameras without active illumination are usually mounted close to the front roof module to capture

both driver and passenger (over 75 %), while modules that serve to warn the driver about rear seat occupants are installed directly above the second row. Rear-row monitoring sensors are mostly radar- or ultrasonic-based.

Gesture control cameras are usually located on the dashboard or roof module between the two front seats to enable access by both driver and passenger. Face recognition cameras, similarly to DMCs, are installed so they focus on the driver's face. Some OEMs use the same camera for both driver monitoring and face recognition.

Which part play 3D sensors for in-cabin monitoring?

For now, 2D sensors are prevalent among in-cabin monitoring modules. While a 3D video feed provides more information, the additional hardware requirements imply a higher cost (see Figure 18), and it seems that for most OEMs, 2D data is sufficient for all use cases of occupant monitoring, especially for non-safety-related features.

There are currently two ways to create a 3D video feed that we have seen in vehicles on the market: using a standard stereo camera, or using a time of flight (ToF) camera. In our breakdown of interior sensors in Figure 15, 3D cameras have a share of 4 %: 2 % of ToF cameras and 2 % of stereo NIR cameras.

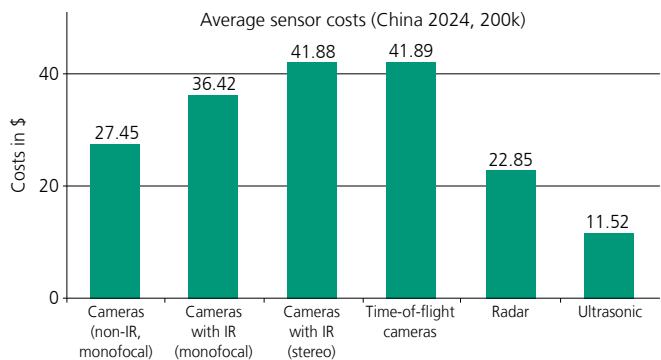


Figure 18: Average cost of interior sensors

The main argument against the usage of 3D sensors is, without a doubt, their costs. As seen in Figure 18, ToF cameras are on average the most expensive modules. For stereo NIR cameras, the costs are on average 15 % higher than for a regular NIR camera, which places it in the same price range as a mono ToF camera.

To sum it up, the trend of 3D data is still a very new one among automotive OEMs. Only one OEM uses stereo driver monitoring cameras, and only two vehicles in our database have ToF cameras – one for gesture recognition and one for driver monitoring.

We are curious to see whether 3D sensors will become more widespread in the coming years, similar to the emergence of Lidar sensors in the last few years.

Thank you, Elena for this analysis of A2MAC1 database!

4.3 Comparison of sensors

The following tables were rated by Fraunhofer IOSB experts on a 3-point scale as highly (green), average (orange) or little (red) suitable sensor performance for the feature detection. In particular, the quality that a sensor can achieve was assessed. Tables 3, 4 and 5 compare the RGB (red, green, blue), NIR (Near infra-

red), FIR (far infrared) and ToF (Time of Fly) sensors. The experts have assessed the quality that a sensor can achieve within its principle boundaries. This means that for example daylight is assumed for RGB and visibility of all relevant objects is assumed – which depends on the camera position. The evaluation is based on the best possible placement, with little or no occlusion, in order to achieve the best possible sensor quality.

Table 3 shows how good the sensor can detect the features.

Detected features	RGB	NIR	FIR	ToF
Object detection	Green	Green	Red	Green
Object localization	Yellow	Yellow	Red	Green
2D Body pose detection	Green	Green	Red	Green
3D Body pose detection	Yellow	Yellow	Red	Green
Body weight detection	Red	Red	Red	Yellow
3D scene reconstruction	Red	Red	Red	Green
Movements e.g. steering wheel rotation	Green	Green	Red	Green
Distance measurement	Red	Red	Red	Green
Face recognition	Green	Green	Red	*
Eyetracking	Yellow	Green	Red	*
Heart rate	Yellow	Red	Yellow	Red
Breath rate	Red	Red	Yellow	Yellow
Body temperature	Red	Red	Green	Red
Visual focus of attention	Green	Green	Red	Yellow

*at DMS position

Table 3: Detected features of sensors

Table 4 shows how good the technical suitability of the sensors are.

Technical suitability	RGB	NIR	FIR	ToF
Dealing with concealment	Red	Red	Red	Red
Function in low light conditions	Red	Green	Green	Green
Function in changing light conditions	Yellow	Yellow	Green	Green
Resolution	Green	Green	Red	Yellow
Temporal resolution	Green	Green	Red	Yellow
Assembly space	Green	Green	Yellow	Yellow
Cost efficiency	Green	Green	Red	Yellow

Table 4: Technical suitability of sensors

Table 5 shows how suitable the sensors are for the different applications.

	RGB	NIR	FIR	ToF
Child Presence Detection	Green	Green	Yellow	Green
Seat belt warning	Green	Green	Red	Green
AirBag control	Yellow	Yellow	Red	Green
Driver state detection	Green	Green	Red	Yellow
Accident recording	Green	Green	Red	Green
Theft protection	Green	Green	Red	Green
Posture corrections	Green	Green	Red	Green
Sleep recognition	Green	Green	Red	Green
Video calls	Green	Yellow	Red	Red
Gesture control	Yellow	Yellow	Red	Green
Activity recognition	Green	Green	Red	Green

Table 5: Applications and suitable sensors

The rating for RGB assumes daylight conditions and good visibility.

The rating for NIR assumes good visibility.

The rating for FIR assumes parked car and interior temperature not equal to body temperature.

The rating for ToF assumes good visibility and high resolution of the drivers face or ToF with a wide field of view. ToF includes a NIR image, which was assumed to be used as well.

4.4 Interview with Prof. Dr. Rainer Stiefelhagen

Prof. Dr. Rainer Stiefelhagen holds the professorship “Computer Science Systems for Visually Impaired Students” at the Faculty of Computer Science at the Karlsruhe Institute of Technology (KIT). He heads the research area “Machine Vision for Human-Machine Interaction” at the Institute of Anthropomatics and Robotics, as well as the KIT’s Center for Digital Accessibility and Assistive Technology – ACCESS@KIT. He received his doctorate in computer science from the University of Karlsruhe (TH) in 2002 and habilitated in 2009.

His research interests include image and video understanding, multimodal interfaces, applications in medical image analysis, assistive technology for seeing impaired users, driver assistance, robotics and surveillance.



Figure 19: Prof. Dr. Rainer Stiefelhagen

What kinds of applications for in-cabin cameras have you been working on?

I have been working together with PhD students on different use cases for in-cabin monitoring. A main focus was activity recognition of all passengers, including the driver. Knowing the historical and current activities allows human machine interactions that respect the context of the user.

Automated cars also promise important mobility improvements for blind passengers. Blind persons can be guided to find objects, doors, seatbelts and can be informed about the general situation by camera-based assistance systems – also in cars.

Which cameras are used in these applications?

We have been working a lot with 3D information for activity recognition and also for assistance of blind persons. A 3D representation of the scene provides independence from changing camera positions. It also measures distance and hence provides a more accurate positioning of body parts and objects. 2D information can only guess distances which means extra efforts. However, it is very good to classify surface types and objects.

We used Time-of-Flight cameras and multi-view cameras in a setup of 3 or more mono cameras. Time-of-Flight cameras require less installation and calibration efforts, while multi-camera setups cover larger areas and suffer less from occlusions. I think the detection area inside the jagged car interior is of very high relevance to get a complete understanding of the situation. E.g. children or objects in the footwell cannot be seen from a camera in the windshield. Of course, it depends on the use case – but my guess is that 3D information and a large coverage is a main feature for in-cabin monitoring. This can be achieved by multi-view systems. Furthermore, they provide additional RGB or FIR signals.

What trends do you see from your research for in-cabin monitoring systems?

Current developments of generative artificial intelligence with powerful foundation models open a large space for new approaches.

Models should be able to handle different but modular sensor input for a holistic representation. Sensors should learn from each other. For this reason, we started a new PhD to investigate multimodal representation learning for in-cabin semantic situation analysis.

Thank you, Prof. Stiefelhagen for your insights!

4.5 Interview with Prof. Dr. Beyerer

Jürgen Beyerer is a professor at the Faculty of Computer Science at the Karlsruhe Institute of Technology. He is also the managing director of the Fraunhofer Institute for Optronics, System Technologies, and Image Exploitation IOSB. He teaches and publishes in the field of computer vision and received his PhD in Engineering with a topic in image processing habilitated in measurement technology. His research interest covers automated visual inspection and image processing and advanced signal processing, environment modeling for intelligent technical systems and human machine interaction. Prof. Dr. Beyerer supervises research in the field of visual perceptual user interfaces and driver assistant systems and consults academic and industry scientist on computer vision measurement technology, environment representation and sensor fusion.



Figure 20: Prof. Dr. Jürgen Beyerer

What is the benefit of using optical sensors in the vehicle interior?

Optical Sensors for different wavelengths provide information about depth, surface reflectance and temperature. This information is unique and not available in cars without cameras. 3D methods like ToF, stereo systems, or multi-camera systems with triangulation provide the shape or geometry of the scene. 2D methods, with or without active illumination like RGB, NIR, enable deductions about the scene's surface reflectance. And thermal imaging sensors like FIR measure surface temperatures. Such information can be helpful for a large variety of safety and comfort functions in future vehicles of private, public and commercial use.

How do you assess the estimation of depth data by monocular cameras?

Cameras in stereo and multi-camera systems can measure depth data accurately and reliably based on triangulation, which is a geometric principle. Stereo systems, similar to human binocular vision, perform a passive triangulation to create a depth map of the scene. With more cameras the principle of triangulation achieves even higher 3D depth accuracy and robustness to occlusion.

Time-of-Flight cameras are based on measuring the travel time of light to an object and back to the sensor. Knowing the speed of light, distances to the object can be calculated. Due to the very short time differences the spatial measurement uncertainty is greater than that of triangulation-based measurements, but is still small enough for vehicle interior surveillance tasks.

Both principles, triangulation as well as time-of-flight, measure depths almost directly.

However, pure 2D images from a single monocular camera can't measure depth directly. However, there are methods to derive depth information from a single 2D image. These are indirect estimations which require extra computing capacity. They are relying on cues similar to how humans estimate depth with one eye – relatively accurate but prone to errors. Known object sizes in the vehicle interior allow for dimension embodiment techniques, and stereo from motion, combining stereoscopy and motion analysis principles. Moreover, end-to-end neural networks and hybrid approaches estimate depth from 2D images. All these methods involve indirect derivations and estimations, with associated error susceptibility, and their adequacy depends on the applications, the required measurement quality and the available computing capacity.

For reliable depth measures, an adequate 3D sensing setup, based on triangulation or time-of-flight, is in most cases the first choice.

What role will camera arrays play in the future?

Camera arrays will become increasingly important. They can, depending on the setup, increase the field of view coverage area compared to single cameras. And if they cover the same area, they can create 3D data via triangulation, crucial for applications needing object distance or absolute 3D position information. Camera arrays can comprise different sensor types (e.g. NIR + ToF, NIR + Thermal) combining their strengths and alleviating their weaknesses, also merge information from various sources for a more accurate reality representation. Combining different optical sensor advantages is a key focus.

What is the future of visual representation in vehicle interiors?

An optimal representation of vehicle interiors and occupants

should describe all relevant aspects and their relationships. An object-oriented world model of the vehicle interior, with abstract representations of humans can achieve this requirement. Such a model has inevitable gaps due to sensor limitations and abstraction. Necessary features, their temporal and spatial resolution, and capture quality must be specified. Scientific institutions like Fraunhofer can contribute models and architectures to this endeavour. Such models also enable simulation and prediction of interior changes, critical for applications like airbags, predicting body movements during crashes. The goal is to develop a cyber-physical model for vehicle interiors to respond to inquiries about past, present, and future states.

Fraunhofer IOSB researchers are also working with the latest neural network-based methods to answer these questions. Large Visual Foundation Models are pivotal, and current AI models are being tested and tuned for interior applications at Fraunhofer IOSB, combining current measuring methods with generative, transformer-driven AI capabilities.

What contribution offers Fraunhofer IOSB to the supply chain of interior monitoring systems?

As an institution for applied research, Fraunhofer IOSB focuses on systems that can reach production readiness in vehicles within 3–5 years. We focus on technologies that are ready for application – or close before. Our contribution includes testing methods, developing methods, implementing proof-of-concepts, and sharing knowledge through publications, consulting and development with clients.

Our research has a long-term foundation and we are a reliable partner for our clients. Long term research projects enable us to dive deep into technologies. Our excellent laboratories are always up to date. A Level 3 automated Mercedes EQS data collection vehicle for public roads, a driving simulator with a mid-sized Audi A3 chassis, and a portable interior monitoring environment are equipped with a variety of cameras and sensors. This comes with a still growing data-base of in-cabin monitoring data, sufficient computing power and the Fraunhofer IOSB Advanced Occupant Monitoring System for research and demonstration.

Thank you, Prof. Beyerer for this insightful interview!

4.6 2D vs. 3D sensing

2D images and 3D images are two different types of visual representations that are used in different contexts and for different purposes.

2D images represent the world as on the human retina and can be interpreted by passengers. Algorithms also interpret this image. 2D images lack depth perception, meaning that they only provide a flat, two-dimensional view of a scene or object. In contrast, 3D images provide depth perception, allowing to see the object or scene in three dimensions, with height, width, and depth, incl. the distance of each pixel from any defined other pixel or view-point, however this 3D image is not interpretable by the passengers.

Most applications in in-vehicles sensing use 2D sensors. Common are NIR and RGB sensors or a combination of both. Also, 3D sensors are used in some vehicles. Stereo vision with two mono-cameras in a defined distance from each other can be used, as well as multi-camera setups with defined positions. More common for car interior monitoring are ToF sensors, that measure the distance between a sensor and an object, based on the phase shift between the emission of a signal and its return to the sensor, after being reflected by an object. This 3D image provides 3D coordinates of each point in the scene.

	2D (e.g. NIR, RGB)	3D (e.g. ToF)
Pro	<ul style="list-style-type: none"> High resolution allows a wide field of view with sufficient resolution High resolution at low cost Better availability of training data Algorithms are more mature and integrated in production cars Provides an image that humans can interpret (monochrome or RGB) 	<ul style="list-style-type: none"> Measures depth information (instead of estimating it) Provides two image modalities (depth and NIR) Allows to reason in space, e.g. distance between objects, distance between passengers and interior Easier segmentation of objects and passengers ToF data is less influenced by the environment and by the texture of objects, which can reduce training data and improve generalization
Contra	<ul style="list-style-type: none"> Lacks depth information and distances between camera and object and between pixels 3D information can only be estimated using machine learning Segmentation of objects and passengers only by estimation with machine learning 	<ul style="list-style-type: none"> Lower resolution of the NIR image, compared to dedicated 2D Higher cost compared to conventional 2D Algorithms for in-cabin sensing have been demonstrated but are less common in production cars

Table 6: Pros and Cons of 2D and 3D interior monitoring sensors

4.7 The dependency of camera position, lens opening angle and resolution

All optical sensors are combined with optical lenses. The lens' opening angle defines the field of view for the sensor. The wider the field of view, the larger the covered area but the lower the pixel resolution per covered inch. The larger the distance from the lens, the smaller the pixel resolution and hence there are less pixels to identify objects and passengers in the back row when the camera is mounted at the windshield. In consequence, the position of the camera in the vehicle is of paramount importance for a given application. It defines the distance to the object of interest and, hence defines, in dependence of the lens' opening angle and the sensor native

resolution, the available pixel resolution for the application. The position also defines the viewing angle of the object of interest, e.g. if the driver's face is visible from front, side or not visible at all.

The position further defines which parts of the cabin interior are occluded by the vehicle's geometry or by passengers or by hands, arms or objects moving around in the cabin. This is often the case when a hand is placed on the steering wheel right in-between a camera positioned behind the wheel and the driver's eyes. In this moment the camera cannot see the eyes and cannot detect sleepiness or distraction. Because of this paramount importance of camera position for field of view and available pixel resolution, we have exemplarily

conducted a proof-of-concept study with an automotive ToF sensor provided by Infineon, mounted in a camera with a 120° lens. The results reveal suitability of the sensor for passenger monitoring and in-cabin monitoring, with sufficient resolution for the front row and limiting resolution for the back row. Driver monitoring applications like distraction detection or face authentication would require a different lens opening angle and different position in the car. Also, a different position, e.g. in the roof above the back row, allows to cover the back row with a higher resolution and less occlusion by the front seats.

Because of these dependencies and relevance of the camera position, a principal differentiation is, if the camera needs to cover as much of the cabin interior as possible (wide angle lens, high resolution sensor, positions with little occlusions, fusion of different cameras) or if the camera shall monitor a very specific field, e.g. the drivers face (narrow angle lens, position without occlusions).

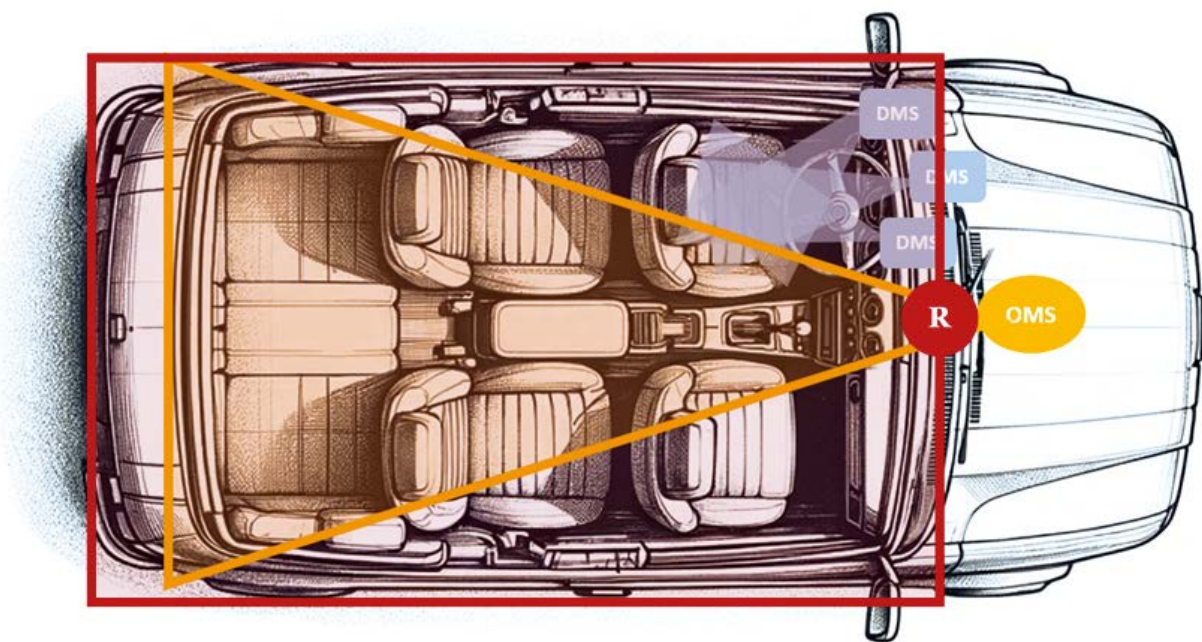


Figure 21: Typical camera positions with different lens opening angles

R: Radar Sensor

DMS: Driver Monitoring Systems (2D or 3D), 40–65°, Driver Focus, e.g. with ToF

OMS: Occupant Monitoring System (2D or 3D), 100–120°, e.g. with ToF

4.8 Evaluation of ToF for in-cabin monitoring

Time-of-Flight (ToF) sensors in the vehicle interior hold significant potential for monitoring purposes due to their robustness against environmental factors. This chapter provides a comprehensive overview of the results obtained from an evaluation project examining the application of Infineon's ToF sensors for vehicle in-cabin monitoring. It delves into both the technical specifications of the Infineon ToF sensors and their practical implementation and assessment within the project.

An interview with Martin Lass, Senior Manager of Product Marketing at Infineon Technologies AG, offers valuable insights into the importance of depth measurements for vehicle monitoring and the necessary specifications for safety applications. The summary of the project results and the interview presents an in-depth analytical examination and interpretation of the performance of ToF sensors in the vehicle interior, along with their potential applications and challenges.

Project Results/Summary

Fraunhofer IOSB test of Infineon ToF:

The **position of the camera** in the car is most important for the coverage and hence for the supported applications.

Additionally, the opening angle of the lens defines the coverage area and resolution of focus areas. A $\sim 120^\circ$ horizontal angle provides the best compromise of large coverage area and good enough resolution. Typical detection functions of ToF are possible with VGA resolution and 120° lens:

- Body pose and limbs distance measuring
- Object detection and distance
- Position of adjustable elements (seat, steering wheel etc.)

This allows reliable applications such as:

- Passenger count and position
- Passenger body pose and activities
- Objects recognition and position

Narrow lenses increase resolution of relevant areas such as driver's face and supports:

- Person authentication
- Eye lid opening and gaze area segmentation

Infineon Sensor Specs

- REAL3™ ToF imager IRS2877A(S)
- VGA system resolution (640 x 480 pixels)
- 4 mm image circle enabling small lens sizes
- Optimized for 940 nm IR light with up to 100 MHz modulation frequency
- Flexible imager operation after each frame: adaptation of frame rate, exposure time or modulation frequency
- Suppression of Background Illumination (SBI) in every pixel for best dynamic range
- Optical PG-LFBGA-65 package, 9 x 9 mm
- ISO 26262 (ASIL-B) compliant

Martin Lass

Senior Manager Product Marketing
3D Time-of-Flight Imager
Infineon Technologies AG



Figure 22: Martin Lass, Infineon

Interview

Why is depth measurement relevant for in-cabin monitoring?

"The real distance measurement of a ToF-sensor in combination with the simultaneous available grey scale image is the most powerful data of the environment. This enables unique use-cases like spoofing robust face authentication and innovative HMI interaction where seamless user-experience is key. A direct, in real time measured 3D body model is also essential for reliable occupant classification and smart restraint and airbag systems."

What specification is required for safety use cases?

"A VGA depth resolution meets the sweet spot between required pixels for the requested detections, calculation efficiency, cost and size. Furthermore, an ISO 26262 compliant imager incorporating on-chip safety mechanisms is mandatory for any safety design."

Time of flight (ToF) sensors hold high potential for in-cabin monitoring. The advantage of ToF sensors is their robustness towards environmental influences like lighting conditions and shadows. In addition, they provide two different image modalities in the form of a Near Infrared (NIR) image as well as a depth image. One drawback specific to ToF is the limited resolution compared to 2D cameras. Disadvantages they share with all cameras relate to the coverage area and possible occlusions depending on the position and interior landscape. In an evaluation project of an Infineon REAL3™ ToF Sensor,

we estimated possible coverage and detection capabilities for relevant interior features with a $120^\circ \times 90^\circ$ lens. A $\sim 120^\circ$ lens covers most of the interior, depending on the camera position. A narrow lens should be used to increase resolution in specific areas of interest, such as the driver's face.

For the project, we simulated the Infineon ToF sensor in the Fraunhofer IOSB Audi A3 driving simulator cabin in different mounting positions. To evaluate these in the context of in-cabin monitoring, we simulated data according to a ToF-camera

specification provided by Infineon. This camera has a field of view of 120° x 90° and VGA system resolution (640 x 480 pixels).

The data for the evaluation was simulated in the 3D animation tool Blender using a to scale model of an Audi A3. On the front seats we placed two people of different genders at the borders of the statistics for body height. The small female driver was rendered in two different poses either holding a smartphone to her ear, like calling someone, or holding the smartphone in front of her for texting. The co-driver is a large male. He is rendered in a bent over position to interact with the infotainment system.

This lets us evaluate a typical driver’s position doing common secondary tasks as well as an out-of-position scenario of the co-driver that causes additional occlusions, depending on the camera viewpoint.

We rendered this scene from five different camera locations from both A-columns at the top and bottom end, as well as above the interior mirror. All these camera locations are viable for in-cabin monitoring, but they pose different challenges. The following discussion of the rendered images focusses on the viability of the introduced ToF sensor at the respective position including an expert opinion on the overall field of view and occlusions, body pose estimation, smartphone detection, head tracking and eye tracking.

The basis of this evaluation were manual measurements of different attributes of the passengers, like their overall bounding box, face bounding box, eye corner distance and bounding box around a smartphone (see Figure 22).

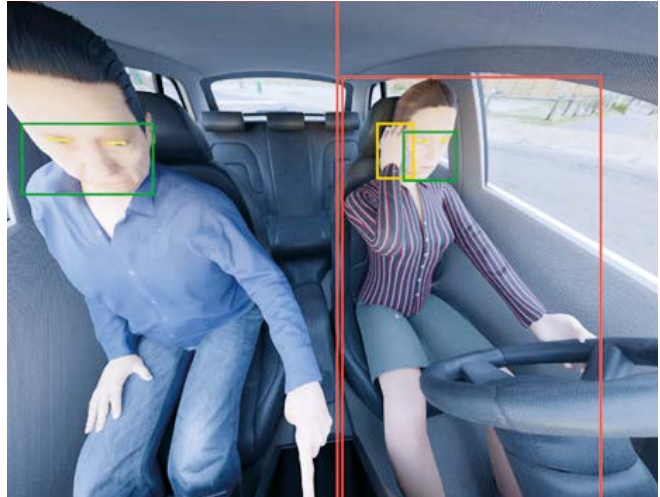


Figure 23: Overview of the measured features for the evaluation

In the following, the five different camera mounting positions (see Figure 23) are discussed in detail and the corresponding feature measurements (in pixels) are summarized in Table 7.

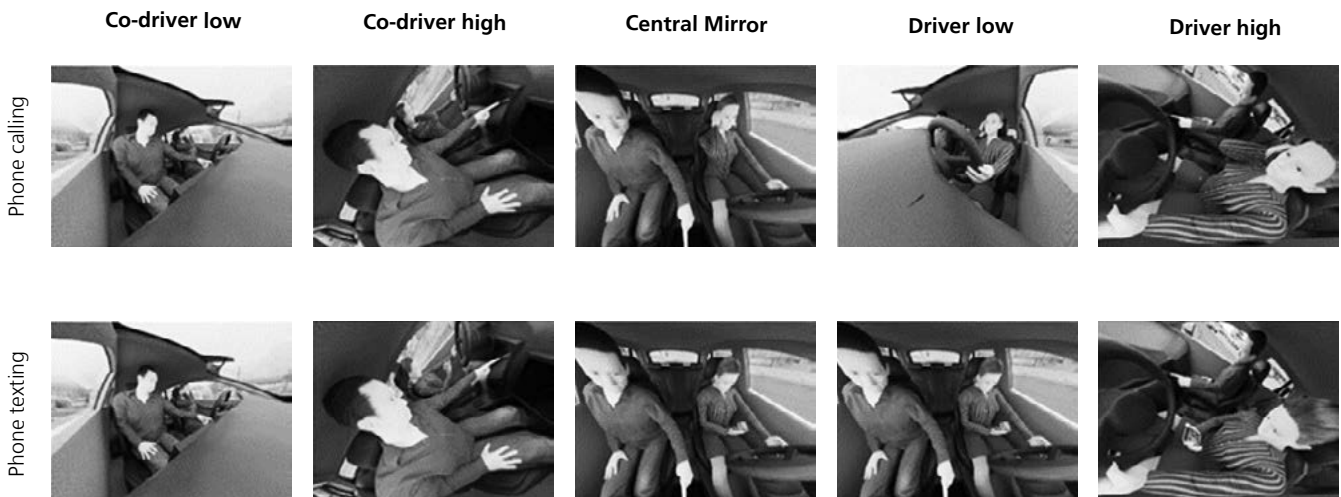


Figure 24: Simulated sensor data for five different camera positions and two different secondary activities of the driver

frame	driver bbox	Co-driver bbox	driver face bbox	Co-driver face bbox	driver right eye corner distance	driver left eye corner distance	smartphone lap bbox	smartphone car bbox	lap eye right	lap eye left
Co-driver low	120 x 136	283 x 300	18 x 21	32 x 48	4	2	9 x 10	24 x 20	2	-
Co-driver high	256 x 160	608 x 376	30 x 30	80 x 88	3	-	11 x 33	36 x 33	2	-
Central Mirror	240 x 410	317 x 480	46 x 46	109 x 96	9	7	56 x 18	32 x 48	6	5
driver low	176 x 272	53 x 120	34 x 37	-	6	8	-	8 x 22	5	7
driver high	640 x 264	259 x 198	120 x 92	33 x 50	15	21	54 x 51	57 x 22	12	16

Table 7: Comparison of different features of the driver and co-driver for different camera positions with Infineon VGA resolution and 120° x 90° lens

All measurements are in pixels, either a line of pixels or bounding box (bbox) square.

Green indicates good suitability. Orange indicates limited suitability because of resolution or occlusion. Red indicates unsuitable.

Co-driver low A-column position: Compared to the driver side this position on the co-driver side is much more usable because there is no steering wheel to block the view. However, the driver's resolution in the image is still low, and occlusions by the co-driver and the challenging side view of the driver make this position difficult for monitoring the driver.

Co-driver high A-column position: This camera position produces a mirror image compared to the driver side. All advantages discussed for that view apply to the co-driver instead of the driver. However, monitoring the co-driver in greater detail than the driver is usually not advantageous. In addition, depending on the posture, the co-driver occludes the driver severely.

Central mirror position: This is the camera position favored by many for in-cabin monitoring. It provides an equal view of the driver and the co-driver. Depending on the slope of the front screen and the dimensions of the vehicle interior, the field of view of the lens limits what can be captured. By angling the camera slightly towards the driver, it is possible to monitor this seat more closely. Overall, all features are well visible but for detailed eye tracking the resolution is too low. However, it should be possible to determine if the passengers look left or right. In general, this view limits occlusions and interference between the passengers. The smartphone is also well visible for both body postures of the driver.

Driver low A-column position: In this view the steering wheel blocks the view severely. This position only works for

monitoring the driver's head. Such a driver monitoring use case, a lens with a narrow field of view should be used.

Driver high A-column position: This camera position is well suited to monitor the driver with higher detail than the co-driver. It offers the best resolution for the driver's face and eyes. Although depending on the driver's position the driver's face is only visible from the left side obscuring the right eye. While monitoring the driver from above works well to detect the smartphone when texting, it restricts the use of eye tracking because eyelashes and deep-seated eyes can obscure the pupil of the eye.

Applications

With a 120° lens, most of the interior can be covered with only one camera. The best position for most applications appears to be close to the rearview mirror in the windshield. This position provides good enough resolution for most applications and even covers some parts of the backseat and front passengers. It enables 3D.

- Body pose and limbs distance measuring
- Object detection and distance
- Position of adjustable elements (seat, steering wheel etc.)

This allows reliable applications such as:

- Passenger count
- Passenger position
- Passenger body pose and activities
- Hands close to wheel detection
- Objects recognition and position

For driver monitoring applications a narrow field of view lens of about 50° is recommendable for the ToF. This increases the 3D pixel (voxel) resolution of the driver's face and upper body which allows applications similar to those known from smartphones with ToF sensors:

- Secure Person authentication
- Eye lid opening angle for sleepiness detection
- Robust eye gaze and head tracking for driver distraction measurements and interaction by eye-tracking

Usually, the high resolution requirement for wide field of view cameras are regarding the 2D information and not on the depth resolution. As such, a ToF camera can also be combined with other cameras, such as RGB or RGB-NIR.

Summary

Overall, this study shows that the simulated ToF camera can work well for driver and passenger monitoring in 3D. ToF provides measured depth information together with a 2D NIR image. The evaluated camera setting provides a good overview when positioned at the central mirror. Depending on the model of the car the field of view may be a bit larger or narrower. The camera positions on the driver side improve the resolution

for monitoring the driver while still allowing to detect rough features of the co-driver. 3D Body pose estimation is possible from all camera positions with this setup. Occlusions are a challenge for certain camera positions. Head tracking is possible with this camera setting from most positions. Side views are more challenging for this task. For precise eye tracking the resolution is too low in general. For some camera positions it should be possible to roughly determine if the passengers' eyes are looking straight, left or right. Regarding the detection of objects, it is advantageous if the camera is viewing the interior from above because there are less occlusions. Object detection capability appears to be good with the given resolution for objects of the size of a smartphone or larger for the front seats. For person authentication and for measuring the eyelid opening angle, we recommend a narrower field of view, of e.g. 50°. The 3D information allows measuring the position in the interior and a better size estimation. 3D information is also very helpful for positioning body parts, tracking the body pose and doing accurate size and weight estimate of the occupants. This is particularly useful for airbag and seat adjustments and activity recognition.

5. State of the Art in Sensors, Training Data and Algorithms

The state of the art in in-cabin monitoring with cameras is rapidly advancing, driven by technological innovation, user expectations and an increasing demand for safety. Some of the key developments in the field are:

- **Automotive-grade camera technology:** The latest camera technology includes automotive-grade high-resolution sensors, depth sensors and temperature sensors with advanced architectures and a well-balanced cost-benefit ratio. This allows OEM now to specify and select from various suppliers the adequate hardware. Single-sensor as well as multi-sensor setups can be realized from the hardware side.
- **Advanced image processing:** Software for image processing and computer vision is advancing rapidly. Limitations are reasonable computing power and energy consumption, data traffic – for higher level computer vision, training data. This implies that significant efforts are still required for implementation of sensors for a specific application in a specific vehicle. However, many demonstrations have shown proof-of-concepts and some singular applications are already implemented in series production vehicles. However most applications cannot yet be bought from the shelf and require thorough sensor selection, specifi-

cation and image processing development efforts. Even applications offered from shelf still require quality testing and may fail in meeting the requirements. Most DMS safety applications in series production still do not meet safety and user experience requirements. In a 2024 test, the Insurance Institute for Highway Safety (IIHS) rated non of the DMS with the label "good" (see IIHS safeguard ratings for partial automation) [20].

- **Sensor fusion:** Sensor fusion is likely to propel the quality and quantity of in-cabin monitoring to a new and finally really useful level. Sensor fusion in in-cabin sensing is however not very advanced yet. Most activities are in-between research and development. Fraunhofer IOSB and other research organizations conduct PhD thesis and develop prototypes for certain applications. Successful demonstrations have been seen in fusing NIR and depth images from ToF sensors. The most promising target appears to be a combination of ToF sensors incl. NIR and depth images with high resolution RGB cameras.
- **Applications and training data:** One of the main outcomes of this paper is the conclusion, that the number of possible applications for in-cabin monitoring is enormous and there are possibly even more applications to be discovered in

the future. Today state-of-the art in algorithms require training for each application, and finetuning for each vehicle. The number of possible applications for series production is mainly limited due to the limitation of available training data and the effort to finetune models for each application. There are several limiting factors for the availability of training data:

- Training data collection, data processing and annotation requires a lot of effort. Only few data sets from universities are available [22] [23]. Most datasets from companies are closed and not shared.
- Most applications require processing of personal data in the training. Or training data needs to be anonymized which increases efforts and costs. Legislations such as GDPR protect users, but they also increase the cost for model development. Many applications (alone) do not promise a respective return of invest.
- **Multi-modal visual foundation models:** Very new advancements in large vision models (similar to large language models), promise a revolution of image processing, based on large image foundation models. Even 2024 models already interpret situations for in-cabin monitoring (see Figure 24). However, accuracy and relevance are still under investigation in research projects, such as "KARLI" and "SALSA" Projects by Fraunhofer IOSB [24].

Overall, the state of the art in in-cabin monitoring with cameras is characterized by advanced camera technology, the need to finetune AI and ML algorithms to the applications, and significant advancements in foundation models for computer vision. These developments are driving the increase of in-cabin monitoring systems.

Figure 25: ChatGPT4 in April, 2024, alerting car drivers based on an in-cabin image

Prompt: You are a personal driving assistant system that politely warns and educates novice drivers. How would you, in short sentences, warn the car driver in this image about an upcoming hazard on the road? Describe what you see in the image. Personalize your answer to the driver activity and the whole scene in the image.

ChatGPT about the scene: In the image, there are two gentlemen seated inside a vehicle, smiling and looking towards the camera, indicating they are posing for a selfie. The driver, on the right, is wearing a sweater over a collared shirt and the passenger on the left is in a casual t-shirt. They seem to be in a good mood, sharing a moment of joy. The driver has one hand on the steering wheel and the other hand making a fist bump gesture with the passenger.

ChatGPT alert to the occupants: It's great to see you both enjoying the drive, but let's keep safety in mind. Driver, please keep both hands on the steering wheel at all times while the vehicle is in motion. It helps maintain control and react quickly if needed. Also, make sure your full attention is on the road and not the camera. Distractions, even for a second, can increase the risk of an accident. Always keep your eyes ahead and be mindful of your surroundings. Have a safe and pleasant journey!

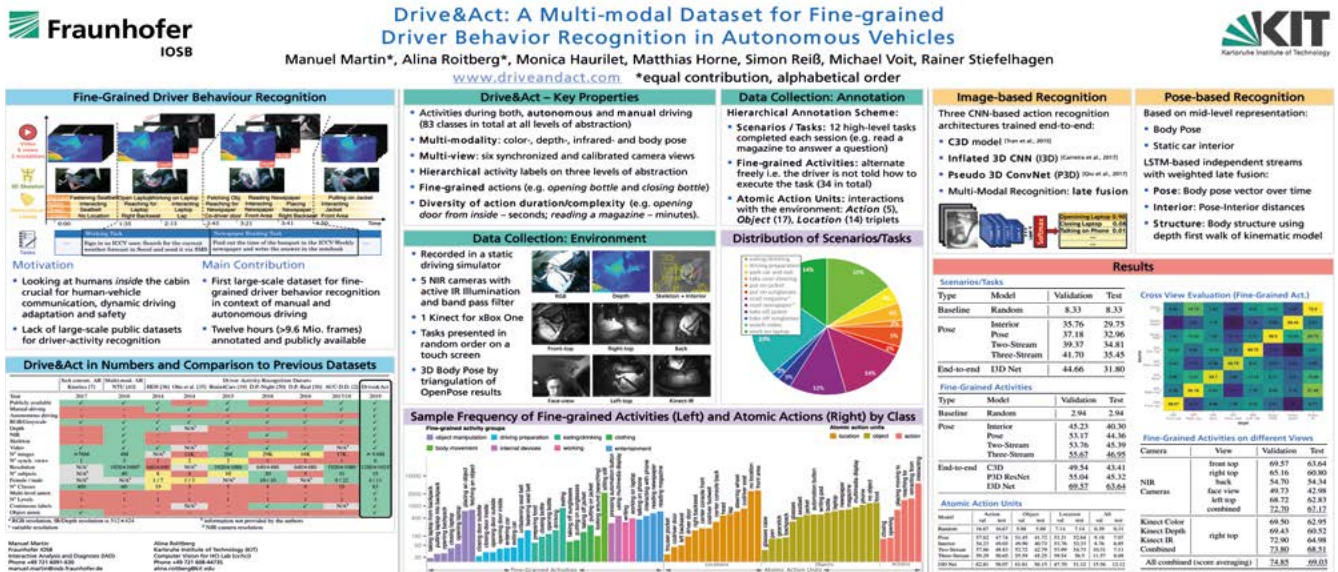


Figure 26: Drive & Act dataset for driver activity recognition training

6. Conclusions

Camera based in-cabin monitoring systems are getting increasingly popular and important. Human error (e.g. distraction, sleepiness, impairment) is one of the major contributors to accidents. With automation of the driving task new requirements emerge for driver attention (e.g. hypovigilance, sleep, out-of-position).

Legislation addresses these risks with international initiatives to implement driver monitoring systems. The potential of driver monitoring systems is enormous, due to the high rate of human error. It may be comparable to the introduction of seatbelts and airbags with regards to the safety impacts.

In the slipstream of safety functions, a huge number of applications for camera based in-cabin monitoring is waiting to be developed. Technology providers have demonstrated several applications, user studies reveal a high quantity of possible applications and smart phone and smart home technologies deliver an increasing bunch of applications.

We see the first introduction of in-cabin monitoring applications in the area of safety, followed by entertainment and comfort features. On the horizon, health monitoring applications are emerging, which will require intensive sensor fusion.

Our impression is, that it is difficult to identify “the one killer application” for in-cabin monitoring. It is rather the sheer number of possible applications that underline the importance for OEM to increase efforts for implementation and to demonstrate the courage to implement the relevant hardware.

The largest potential for in-cabin monitoring is only imaginable right now and many more applications will be imagined in the future. Even third-party developers and after market solutions should be considered, which leverage a huge amount of creativity and development forces.

The main advancements are expected in software development and can be implemented “over-the-air” if the respective hardware is on-board.

We foresee at least two more technology leaps (that follow the current application of machine learning). Sensor fusion of camera arrays or even other sensors will significantly increase the accuracy of the detectable features and also the number of reasonable applications.

Large multi-modal visual foundation models are the next technological leap. Their potential is barely becoming visible in 2024 and is likely to improve general understanding of the interior holistic situation dramatically in the following years.

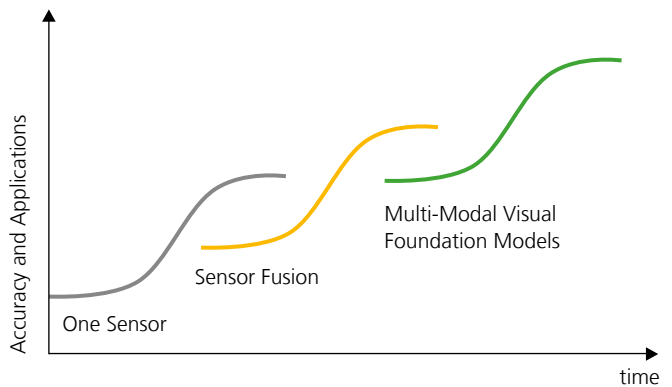


Figure 27: Expected technology leaps in in-cabin monitoring software

Along this technological progress, applications of in-cabin monitoring systems extend beyond safety functions to encompass a wide range of user-centric interactions and assistance mechanisms. From activity recognition to aiding blind passengers, these systems are poised to revolutionize the in-vehicle experience by providing personalized and context-aware services.

The choice of cameras plays a pivotal role in determining the effectiveness and coverage of in-cabin monitoring systems. While RGB cameras offer a high resolution and an image that can be used for video calls and photos, 3D sensors like time-of-flight cameras offer robust depth information. Multi-view camera arrays provide broader coverage and reduced occlusions. The fusion of different sensor modalities, such as RGB with time-of-flight enhances the capabilities of these systems.

As automotive manufacturers strive to integrate in-cabin monitoring systems into production vehicles within a relatively short timeframe, the role of research institutions like Fraunhofer IOSB becomes increasingly crucial in bridging the gap between cutting-edge research and practical implementation.

By leveraging these technologies effectively, automotive manufacturers can create safer, more intuitive, and personalized driving experiences for occupants and thereby shaping the future of mobility.

References

- [1] WHO. 2023. Road traffic injuries. Available: <https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries>. [Accessed 12 January 2024].
- [2] S. Klauer, T. A. Dingus, J. Sudweeks, T. Neale and D. J. Ramsey. 2006. The Impact of Driver Inattention on Near-Crash/Crash Risk: An Analysis Using the 100-Car Naturalistic Driving Study Data. Technical Report, U.S. Department of Transportation, National Highway Traffic Safety Administration.
- [3] National Center for Statistics and Analysis. 2024. Distracted driving in 2022 (Research Note. Report No. DOT HS 813 559)., National Highway Traffic Safety Administration
- [4] European Union. (2019). Regulation (EU) 2019/2144 of the European Parliament and of the Council of November 27, 2019 on type-approval requirements for motor vehicles and their trailers, and systems, components and separate technical units intended for such vehicles, as regards their general safety and the protection of vehicle occupants and vulnerable road users (OJ L 325, December 16, 2019, ELI: <http://data.europa.eu/eli/reg/2019/2144/oj>)
- [5] European Union. (2023) Commission Delegated Regulation (EU) 2023/2590 of July 13, 2023 supplementing Regulation (EU) 2019/2144 of the European Parliament and of the Council by laying down detailed rules concerning the specific test procedures and technical requirements for the type-approval of certain motor vehicles with regard to their advanced driver distraction warning systems and amending that Regulation (OJ L, 2023/2590, ELI: http://data.europa.eu/eli/reg_del/2023/2590/oj)
- [6] European Union. (2021). Commission Delegated Regulation (EU) .../... supplementing Regulation (EU) 2019/2144 of the European Parliament and of the Council by laying down detailed rules concerning the specific test procedures and technical requirements for the type-approval of motor vehicles with regard to their driver drowsiness and attention warning systems and amending Annex II to that Regulation (Ares (2021) 1075107; ELI: https://www.parlament.gv.at/dokument/XXVII/EU/58896/imfname_11061661.pdf).
- [7] H.R.3164 - 117th Congress (2021–2022): Hot Cars Act of 2021. (May 13, 2021). Available: <https://www.congress.gov/bill/117th-congress/house-bill/3164/text>
- [8] S. 1406 - 117th Congress (2021–2022): Stay Aware For Everyone Act of 2021. (April 28, 2021). Available: <https://www.congress.gov/bill/117th-congress/senate-bill/1406>
- [9] H.R.3684 - 117th Congress (2021–2022): Infrastructure Investment and Jobs Act. (November 15, 2021). Available: <https://www.congress.gov/bill/117th-congress/house-bill/3684>.
- [10] Alliance for Automotive Innovation, Available, November 1, 2023, Press Release: Lifesaving Rear Seat Reminder Systems: Now Available in 215+ New Vehicle Models. Available: <https://www.autosinnovate.org/posts/press-release/2023-industry-heatstroke-report>. Chapter 10: Interpretation and Conclusion Confidential draft version
- [11] National Automotive Standardization Technical Committee 2022. Performance requirements and test methods for driver attention monitoring system, GB/T 41797-2022. Available: <https://www.chinesestandard.net/PDF/English.aspx/GBT41797-2022>.
- [12] Benedikt Spannochci. 3demotion. 2024: Regulatory Radar Available: <https://emotion3d.ai/regulatory-radar/>. [Accessed December 13, 2023]
- [13] MLIT. 2020. SIP-adus Workshop 2020: Efforts of Road Transport Bureau, MLIT for the realization of Automated Driving. Available: https://en.sip-adus.go.jp/evt/workshop2020/file/jg/08JG_06E_Tada.pdf.
- [14] MORTH.2022. Finalized Draft Automotive Industry Standard, Driver Drowsiness and Attention Warning Systems for M, N2 and N3 category vehicles. Available: https://morth.nic.in/sites/default/files/ASI/1_Draft_AIS%20184_DF.pdf.
- [15] World Forum for Harmonization of Vehicle Regulations (ECE/TRANS/WP.29/1147).2021. Framework document on automated/autonomous vehicles. Available: https://unece.org/sites/default/files/2022-02/FDAV_Brochure%20-%20Update%20Clean%20Version.pdf.
- [16] United Nations. 2024. New UN regulation paves the way for the roll-out of additional driver assistance systems, February 1, 2024, Available: <https://unece.org/media/Sustainable-Development/press/387961>. [Accessed February 15, 2024]
- [17] Euro NCAP. 2022. Euro NCAP Vision 2030: A Safer Future for Mobility. Available: <https://cdn.euroncap.com/media/74190/euro-ncap-vision-2030.pdf>.
- [18] Euro NCAP. 2023. Test and assessment protocol - child presence detection. Available: <https://www.euroncap.com/media/79888/euro-ncap-cpd-test-and-assessment-protocol-v12.pdf>.
- [19] ANCAP Safety. 2022. Future view on testing & assessment: an expanding focus to 2030. Available: <https://s3.amazonaws.com/cdn.ancap.com.au/app/public/assets/c2f8d124e9d551a7596111601bddd1fa419da9e0/original.pdf?1669765012>.
- [20] IIHS. 2022. IIHS creates safeguard ratings for partial automation. Available: <https://www.iihs.org/news/detail/iihs-creates-safeguard-ratings-for-partial-automation>.
- [21] C-NCAP. 2024. China Automotive Technology and Research Center Co. Ltd. (CATARC). Available: https://www.safetywissen.com/object/B04/B04.f9z738952glgzjm4uyf38699momnm263845405099/safetywissen?trk=article-ssr-frontend-pulse_little-text-block.
- [22] Martin, Manuel; Roitberg, Alina; Haurilet, Monica; Horne, Matthias; Reiß, Simon; Voit, Michael; Stiefelhagen, Rainer. 2019. Drive&Act: A multi-modal dataset for fine-grained driver behavior recognition in autonomous vehicles. Available: https://driveandact.com/publication/2019_iccv_drive_and_act/2019_iccv_drive_and_act_poster.pdf.
- [23] Ortega, Juan Diego; Köse, Neslihan; Rodriguez; Paola Natalia; Chao, Minan; Unnervik, Alexander; Donce, Marcos Nieto; Madurga, Oihana Otaegui; Salgado, Luis. VICOMTECH. 2020. Available: <https://www.vicomtech.org/en/rdi-tangible/publications/publication/dmd-a-largescale-multimodal-driver-monitoring-dataset-for-attention-and-alertness-analysis>. Chapter 10: Interpretation and Conclusion Confidential draft version
- [24] Diederichs, F., Wannemacher, C., Faller, F., Mikolajewski, M., Martin, M., Voit, M., ... & Piechnik, D. (June 2022). Artificial intelligence for adaptive, responsive, and level-compliant interaction in the vehicle of the future (KARLI). In International Conference on Human-Computer Interaction (pp. 164-171). Cham: Springer International Publishing. KARLI-Projekt Website: <https://karli-projekt.de/en/start>

Layout:

Atelier F. Bruns, Karlsruhe, Germany

Image rights:

A2MAC1 (p. 31, 32)

Fraunhofer IOSB (p. 1, 14, 17, 18, 35, 38, 40)

Melexis (p. 16)

private (p. 19, 34, 39, 43)

Printing:

Späth Media, Baden-Baden, Germany

Publication date:

May 22, 2024

© Fraunhofer IOSB, Karlsruhe, 2024

Contact



Dr.-Ing. Dipl.-Psych. Frederik Diederichs
Head of Perceptual User Interfaces
Research Unit Human-AI-Interaction
frederik.diederichs@iosb.fraunhofer.de

Fraunhofer IOSB
Fraunhoferstraße 1
76321 Karlsruhe, Germany