Augmented Vision and Reality 5

Anestis Terzis *Editor*

Handbook of Camera Monitor Systems

The Automotive Mirror-Replacement Technology based on ISO 16505

Springer

Augmented Vision and Reality

Volume 5

Series editors

Riad I. Hammoud, Kokomo, IN, USA
Lawrence B. Wolff, New York, NY, USA

More information about this series at http://www.springer.com/series/8612

Anestis Terzis
Editor

Handbook of Camera Monitor Systems

The Automotive Mirror-Replacement Technology based on ISO 16505

Springer

Editor
Anestis Terzis
Department of Electrical Engineering
and Information Technology
Ulm University of Applied Sciences
Ulm
Germany

ISSN 2190-5916 ISSN 2190-5924 (electronic)
Augmented Vision and Reality
ISBN 978-3-319-29609-8 ISBN 978-3-319-29611-1 (eBook)
DOI 10.1007/978-3-319-29611-1

Library of Congress Control Number: 2016930541

Printed on acid-free paper

This Springer imprint is published by SpringerNature
The registered company is Springer International Publishing AG Switzerland

To my children
Angelos and Maria

Preface

This edited book, *Handbook of Camera Monitor Systems—The Automotive Mirror-Replacement Technology based on ISO 16505*, aims to provide a comprehensive overview of the science and technology of camera monitor systems (CMS). The content ranges from the ISO 16505-based development aspects to practical realization concepts in vehicles. In addition, it serves as a single reference source with contributions from leading international CMS professionals and academic researchers combining technological as well as ergonomic aspects. Most of the authors, including the editor, have been involved in the international standardization and regulation of this technology over the past years.

The replacement of automotive mirrors by CMS is a growing research and development field. Such systems can be used to improve factors in the driver's indirect vision, to improve aerodynamics, and to optimize the fuel economy of new vehicles. The CMS idea has existed for concept cars for decades, but until now there was no international legislation framework for such solutions in series production vehicles. The first milestone was the publication of the international standard ISO 16505 "Road vehicles—Ergonomic and performance aspects of Camera Monitor Systems—Requirements and test procedures" in the year 2015. In combination with the latest version of UN Regulation No. 46, the normative framework of ISO 16505 permits CMS to replace mandatory rear-view mirrors. Working on CMS requires specific knowledge of the technical fundamentals, standardization and regulation aspects, as well as specific automotive requirements and the relevant ergonomics. Although literature exists which covers subtopics, e.g., automotive camera technology, until now no text combining all the required disciplines in one special book dedicated to CMS exists. Furthermore, methods and results for the ergonomic design of such systems are included.

The book is organized into five parts.

Part I "CMS System Design and Standardization and Regulation Aspects" is dedicated to the system design of passenger as well as commercial vehicles. It includes a contribution covering the CMS-specific standardization and regulation aspects. A key topic of CMS, which is addressed by a special contribution, is the resolution and the sharpness of the complete system.

Part II “Fundamentals of Automotive Technology for CMS” covers the relevant fundamentals of automotive imagers, video interface technology, and embedded image processing components. All contributions present the content with regard to CMS. The optical effects in camera monitor systems in combination with optical measurement setups are presented in a special contribution.

Part III “Human Visual Perception and Ergonomic Design” starts by presenting the properties of human visual perception with respect to CMS. It includes contributions covering the ergonomic design of CMS for the very demanding commercial vehicles scenario.

Part IV “CMS Tests and Concepts for Passenger Cars and for Commercial Vehicles” includes a study comparing CMS and conventional exterior mirrors and which made its assessment using test drives and static tests under different external conditions. It is a unique text covering these aspects with respect to the ISO 16505 and UN R.46 requirements. The German Federal Highway Research Institute (BASt) carried out this study on behalf of the German Federal Ministry of Transport and Digital Infrastructure. A contribution with concepts for commercial vehicles is also included in this part of the book.

Part V “Advanced Topics” provides content with direct or indirect relevance to CMS. It begins with a discussion of demanding scenarios in CMS and includes image-quality criteria. A special contribution presents a novel approach for intuitive motion and depth visualization for rear-view camera applications. The book concludes with a dedicated contribution to the very important functional safety aspects of CMS based on ISO 26262. It explains what hazards could arise in the context of CMS and how they can be systematically investigated.

Acknowledgments

I would like to acknowledge all those who helped in creating this book. First and foremost, I want to thank all 23 authors of the book's chapters. Without your willingness to invest your valuable time and experiences, this book would never have seen the light of day in the current form.

I know most of the authors from my time working on CMS standardization and regulation. This work started officially in 2009 when the Informal Group Camera Monitor Systems (IGCMS) was established in order to work on the technical content of United Nations Regulation No. 46. In 2010, the CMS standardization work began within the framework of the International Organization for Standardization (ISO). We started the work as an international expert group at a kick-off meeting in London in 2010. Until 2015, when the ISO 16505 was finally published and the reestablished IGCMS finalized their work, we worked very intensively and constructively on this topic. We had working meetings all around the world including such places as Berlin, Milan, Paris, Stockholm, and cities in Japan and the USA. We were supported by uncountable web-conferences. Every expert enriched the group with their special area of expertise and their individual personality, I learned a lot during that time. It was an honor for me to work together with these people and I want to thank everyone for the great cooperation. Especially the convenors of the ISO WG, the delegates from the ministries, the IGCMS chairman and secretary, the vehicle manufacturers' associations and all representatives from the automotive industry and from the technical services. All of them had an influence on how I see and understand camera monitor systems, and thus, by extension, on editing this book. I also want to thank all my direct project cooperation partners in the automotive industry. It was at one of the last expert meetings where I started thinking about what an enrichment it would be to combine this expert knowledge in a special book and make it available to the technical and scientific society. It was the birth of the book idea.

Ulm University of Applied Sciences has provided a supportive environment and the scientific freedom for me to explore this advanced camera-based driver assistance system and to establish it as a research discipline. Especially, the technical and scientific facilities of the Institute of Automotive Systems Engineering (IFS) and the Institute of Communication Technology (IKT) allow for a system-level approach ranging from chip design and EMC-investigations to real vehicle implementations. My research team supported my work in the last years through student projects and theses both for SW and HW aspects of CMS. I owe a debt of gratitude to all of them.

I also want to thank the German national standards body DIN (Deutsches Institut für Normung e.V.) for giving permission[1] to use material from ISO 16505.

Furthermore, I greatly appreciate Springer for publishing this book and for the expert guidance provided throughout the book production process.

My whole family has always supported me. My wife, Jula, and our two children have given me the freedom of many evenings and weekends to complete this manuscript. I am grateful for their considerable patience and support.

Ulm
November 2015

Anestis Terzis

[1]The following information applies to every ISO 16505 reference in this book:

"Reproduced by permission of DIN Deutsches Institut für Normung e.V. The definitive version for the implementation of this standard is that edition which bears the most recent date of issue, obtainable from Beuth Verlag GmbH, Burggrafenstraße 6, 10787 Berlin, Germany."

Contents

Editor and Contributors

About the Editor

Prof. Dr. Anestis Terzis

is the head of the Automotive Electronics study course at Ulm University of Applied Sciences in Germany and the coordinator of the university's International Electrical Engineering Program (IEEP). Prior to this, he was with Daimler AG for ten years and worked in "Group Research & Mercedes-Benz Cars Development" in the field of future advanced driver assistance systems.

He was born in 1978 in Heidenheim, Germany, and received his diploma degree in Communications Engineering from the Ulm University of Applied Sciences in 2002. He holds a doctoral degree in Electrical Engineering (Dr.-Ing.) from the University of Erlangen-Nuremberg, Germany.

Professor Terzis is one of the authors of the ISO 16505 content and an expert member of the standardization and regulation committees in the field of camera monitor systems. He was also an expert member of the working group ISO/TC 22/SC 17/SC 3 "Road vehicles—Video communication interface for cameras" (ISO 17215).

In his research and development fields, he has published numerous international conference and journal papers. During his studies, he wrote his first book at the age of 23 together with a fellow student. He has filed more than 20 patents in the areas of digital radio transmission and camera-based driver assistance systems.

He left the Daimler AG Research Center in 2012 to become a professor for Digital Systems Design at Ulm University of Applied Sciences. His lectures include digital technology with VHDL and FPGAs, basics of electrical engineering, and automotive systems. His primary research field is advanced camera-based driver assistance systems.

Professor Terzis combines academic and industrial experience. He is the founder and the director of the Steinbeis Transfer Center "DSI—Digital Systems and Innovations," a company in the Steinbeis Network located in Ulm, Germany. This Center offers consulting, as well as courses, prototypes, and measurement technology (www.stw.de/su/1637).

His contact data can be found at www.hs-ulm.de/terzis.

Contributors

Maxim Bierbach Bundesanstalt für Straßenwesen (BASt), Federal Highway Research Institute, Bergisch Gladbach, Germany

Patrik Blomdahl Vehicle Engineering, Volvo Group Trucks Technology, Gothenburg, Sweden

Alexander Georg Bothe Institute of Ergonomics, Darmstadt University of Technology, Darmstadt, Germany

Michael Brading ON Semiconductor, San Jose, CA, USA

Matthias Esser Mercedes-Benz Cars Development, Certification and Regulatory Affairs Active Safety, Daimler AG, Sindelfingen, Germany

Christian Faber Sensor Technology and Image Processing, University of Applied Sciences Landshut, Landshut, Germany

Alexander Frey Bundesanstalt für Straßenwesen (BASt), Federal Highway Research Institute, Bergisch Gladbach, Germany

Jost Gail Bundesanstalt für Straßenwesen (BASt), Federal Highway Research Institute, Bergisch Gladbach, Germany

Rainer Gutzmer HELLA Aglaia Mobile Vision GmbH, Berlin, Germany

Patrick Heinemann Pre-Development Driver Assistance Systems, Audi Electronics Venture GmbH, Gaimersheim, Germany

Heike Hoffmann Bundesanstalt für Straßenwesen (BASt), Federal Highway Research Institute, Bergisch Gladbach, Germany

Bernhard Kaiser Berner & Mattner Systemtechnik GmbH, Munich, Germany

Brian Keelan ON Semiconductor, San Jose, CA, USA

Rainer Krautscheid Bundesanstalt für Straßenwesen (BASt), Federal Highway Research Institute, Bergisch Gladbach, Germany

Mark Leznik Ulm University of Applied Sciences, Ulm, Germany

Christine Lotz-Keens Bundesanstalt für Straßenwesen (BASt), Federal Highway Research Institute, Bergisch Gladbach, Germany

Mark Müller Socionext Europe GmbH, Langen, Germany

Eiji Oba Atsugi Technology Center, Sony Corporation, Atsugi, Japan

Christoph Rößing Department of Computer Graphics, Eberhard Karls University, Tübingen, Germany

Tobias Schmalriede Continental-Corporation AG, Frankfurt am Main, Germany

Eike Andreas Schmidt Bundesanstalt für Straßenwesen (BASt), Federal Highway Research Institute, Bergisch Gladbach, Germany

Anestis Terzis Department of Electrical Engineering and Information Technology, Ulm University of Applied Sciences, Ulm, Germany

Hieu Tran ON Semiconductor, San Jose, CA, USA

Albert Zaindl MAN Truck & Bus AG, Munich, Germany; Institute of Ergonomics, Technical University of Munich, Munich, Germany

Part I
CMS System Design and Standardization and Regulation Aspects

Automotive Mirror-Replacement by Camera Monitor Systems

Anestis Terzis

Abstract This contribution presents a discussion of automotive mirror-replacement by camera monitor systems (CMS). This is a growing research and development field. Such systems can be used to improve factors of the indirect vision of the driver, to improve the aerodynamics and to optimize the fuel economy of new vehicles. Besides the potential benefits of such a system, there are also potential challenges discussed in this contribution. The normative framework of the new standard ISO 16505:2015 in combination with the latest version of UN Regulation No. 46 enables the replacement of mandatory rear-view mirrors by CMS for vehicles in series production. Based on the characteristics of conventional mirrors, a CMS has to be designed for comparable functionality. This includes the real time behavior, the system resolution, and the dimensions of the displayed object and field of view. In this chapter, beginning with system requirements, incorporating additional standards and regulations, different CMS architectures are discussed. The final subsection presents a system design based on ISO 16505:2015 including selected real values and a description of camera, display, and image processing components of a CMS.

Keywords Camera monitor systems · CMS design · ISO 16505 · UN Regulation No. 46 · Mirror-replacement · CMS architectures

List of Abbreviation

AEC	Automotive Electronics Council
ASIL	Automotive Safety Integrity Level
CMOS	Complementary Metal Oxide Semiconductor
CMS	Camera Monitor System
ECE	Economic Commission for Europe
ECU	Electronic Control Unit
EMC	Electromagnetic Compatibility

A. Terzis (✉)
Department of Electrical Engineering and Information Technology, Ulm University of Applied Sciences, Ulm, Germany
e-mail: terzis@hs-ulm.de

A. Terzis (ed.), *Handbook of Camera Monitor Systems*,
Augmented Vision and Reality 5, DOI 10.1007/978-3-319-29611-1_1

FOV	Field of View or Field of Vision
FPGA	Field Programmable Gate Array
GDC	Graphics Display Controller
HDR	High Dynamic Range
HFoV	Horizontal Field of View
I^2C	Inter-Integrated Circuit
ISO	International Organization for Standardization
LCD	Liquid Crystal Display
LVDS	Low Voltage Differential Signal
MTF	Modulation Transfer Function
OEM	Original Equipment Manufacturer
ORP	Ocular Reference Point
SoC	System on Chip
TFT	Thin Film Transistor
UN	United Nations

1 Introduction

Each year, a significant growth in passenger traffic can be observed. This growth is fueled to a great extent by the increase in motorized individual traffic. A positive development is that despite the increase in the number of vehicles, the annual number of traffic deaths in industrial nations such as Germany has been decreasing over the past three decades. This development can be attributed to political transportation measures, improvements in emergency service, and technical measures. The increase in traffic safety, the responsible use of fossil resources as well as the reduction of CO_2 emissions continue to have high social priority. The replacement of the external vehicle mirrors prescribed by law with camera monitor systems (CMS) can contribute to the fulfillment of these goals. The side mirrors can be eliminated through this technical measure, resulting in reduced air resistance for the vehicle. A camera substitutes for the mirrors and records the driving events behind and beside the vehicle and presents the visual information on a suitable display in real time inside the vehicle. Figure 1 shows an example of a passenger car equipped with mandatory Class III and Class I mirrors (top) and an example where these mirrors are substituted by a CMS (bottom). The cameras are mounted outside the vehicle in the position of the conventional mirror in this example.

Such CMS consist of a camera capturing a field of view, forwarding the signal to, e.g., an electronic control unit (ECU) for additional processing and then using a display to visualize the information for the driver. Figure 2 shows a basic concept of such a CMS.

Individual mobility is now perceived as a characteristic of modern society and as the driving force for economic growth. The demographic shift, especially in

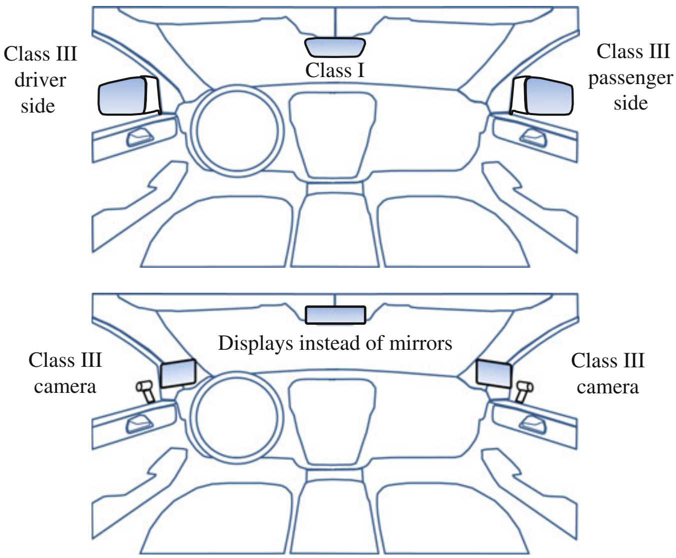


Fig. 1 UN R.46 mirror classes passenger car (*top*) and with CMS example (*bottom*)

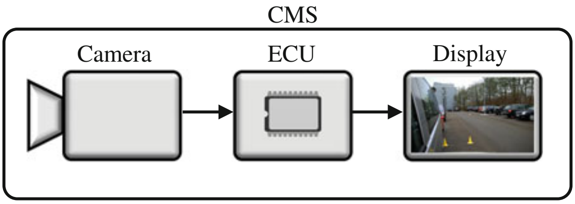


Fig. 2 Camera monitor system

industrial nations will lead to a larger percentage of older drivers in the future. Future vehicles will contain individual innovations for these age groups as well. New systems for vision support such as, e.g., CMS, can contribute in this scenario as well. The innovative vehicle mirror replacement addresses various current social requirements and thus follows the principle of sustainable development in an impressive manner.

1.1 Normative and Regulatory Framework

This subsection presents a brief overview of the normative and regulatory framework of the CMS technology which emphasizes the ISO 16505:2015. A detailed discussion of this topic can be found in this handbook in chapter "Standardization and Vehicle Regulation Aspects of Camera Monitor Systems".

According to United Nations Regulation No. 46 (UN R.46), vehicles must be equipped with systems providing indirect vision [1]. This indirect vision is achieved by using conventional mirrors attached to the vehicle. Based on the new international standard ISO 16505:2015 "Road vehicles—Ergonomic and performance aspects of Camera Monitor Systems—Requirements and test procedures", the technical requirements are outlined for the replacement of such mandatory mirrors by camera monitor systems [2].

The latest amendment to United Nations Regulation No. 46 is based on the ISO 16505:2015 and is the first regulation permitting the use of CMS as an alternative to mirrors for passenger cars as well as for commercial vehicles. The proposal to amend UN R.46 was adopted in November 2015 by the "World Forum for Harmonization of Vehicle Regulations (WP.29)" and is expected to enter into force around August 2016 [3].

Against the backdrop of recent technical developments in modern driver assistance systems, a variety of camera-based systems have been introduced in series production vehicles. The realized functions range from simple rear-view cameras to advanced 360° surround view systems and even night-vision assistance systems including object recognition and sensor fusion. As a rule, such systems offered in series production vehicles are considered to be comfort or assistance systems and cannot be considered as mandatory vehicle systems. ISO 16505:2015, however, focuses on camera monitor systems which are intended to replace systems that are mandatory under current law.

1.1.1 Scope and Structure of ISO 16505

ISO 16505:2015 describes the minimum technical requirements that must be fulfilled by a camera monitor system in order to substitute a mandatory mirror. These requirements are related to safety aspects, ergonomic factors, performance criteria and the testing of such camera monitor systems. The new standard forms the basis of future standard-conforming CMS design and standard-conforming CMS testing carried out by technical service providers. The standard document of ISO 16505:2015 was published in May 2015 and is organized in different chapters, beginning in Chap. 1 with the scope. The philosophy of the standard is to provide a CMS specification on a system-level approach that is independent of different camera, display, and processing technologies. A CMS designer will have different system-related options to design such a system and a vehicle manufacturer will have different options to define the appropriate integration concept into the specific vehicle. Additional functions and systems, such as Advanced Driver Assistance Systems (ADAS), are explicitly stated not to be part of this standard. After the citation of normative references (Chap. 2) the used terms and definitions are described in the Chap. 3 of ISO 16505:2015. There are terms and definitions related to the vehicle and to the mirror and terms and definitions related to the components camera, monitor and to the overall CMS. Finally, fundamental definitions such as modulation transfer functions and luminance aspects are presented. Chapter 3 also includes references to other documents and regulations (e.g., UN R.46 and FMVSS 111 for the USA market) where appropriate. Symbols and abbreviated terms are presented in Chap. 4, followed by general information and use case definitions in Chap. 5, including a list of possible use cases applicable for CMS. The detailed system requirements are presented and discussed in Chap. 6. That chapter begins with the intended use of the system and specifies requirements such as the system availability and field of views (FOV). The phrases “field of view” or “field of vision” are both used and are interchangeable. Specific emphasis is given to parameters such as magnification factors for the different classes, resolution definition, image quality criteria and time behavior of the CMS. In Chap. 7, test methods are defined for verifying whether a system fulfills the specified requirements. Chapters 6 and 7 should be read as a single entity with their content directly linked. A requirement is formulated using the word “shall” and a recommendation by using the word “should”. In Chap. 7, test procedures are described only for the requirements in Chap. 6. For the Chap. 6 recommendations, there is only the information that “Verification is not required”. The main body of the standard ends with Chap. 8, which deals with the functional safety issue. CMS is considered to be safety relevant. To deal with this, only a reference to safety standards, e.g., IS0 26262, is included. There is no description about specific safety relevant parameters because this depends on the very concrete system design, and as already mentioned the philosophy of ISO 16505:2015 is to be as technology independent as possible. A detailed discussion of these very important safety issues can be found in this handbook in chapter “Functional Safety of Camera Monitor Systems”. ISO 16505: 2015 also includes various annexes categorized as normative or informative. Annex A is a normative annex dealing with specific CMS considerations for commercial vehicles. Annex B to G are informative and include content such as formula applications, explanations, guidelines, calculation of parameters and measurement information.

1.1.2 The Way of the Amendment for UN Regulation No. 46

On its own, ISO 16505:2015 does not provide sufficient binding for such mandatory mirror replacement systems. In Germany, for example, the current 2013 version of UN Regulation No. 46 is being applied to systems that enable indirect vision [1]. Vehicles can be given an operating license if they fulfill the minimum technical requirements. The rules laid down under UN Regulation No. 46 came from a commission of the United Nations for Europe (UN-ECE, where ECE stands for Economic Commission for Europe). As part of the UN-ECE, there is a World Forum for the Harmonization of Vehicle Regulations or UN-ECE WP.29. A committee was set up in 2009 to initiate amending UN Regulation No. 46 in order to make camera monitor systems feasible as a replacement for all mirrors. To define the technical parameters and to prepare an amendment, an Informal Group on Camera Monitor Systems (IGCMS) was established. For commercial vehicles it has been allowed since 2005 to replace Class V (close-proximity view device) and VI

(Front-view device) mirrors by CMS. These mirror classes are only used for low-speed tasks such as maneuvering. Due this fact it was not appropriate to use the same requirements for all other mirror classes. One of the results of the work carried out by the IGCMS was a standards initiative for such systems, within the expert framework of the International Organization for Standardization (ISO).

Intensive standardization work was carried out by an international expert group between 2010 and 2014, resulting in the draft of ISO 16505. The IGCMS was reestablished (as ICGCMS II) in 2014. Based on the ISO 16505 draft version, an amendment for UN R. 46 was prepared by the IGCMS II including additional formulations and additional content. It was not possible simply to refer to the complete ISO 16505 standard. As mentioned in Sect. 1.1.1, the ISO contains requirements but also recommendations. The concern was that such recommendations could possibly result in different interpretations of technical services. The IGCMS II work resulted in the proposal of an amendment to UN R.46 that was adopted in November 2015 by WP.29 and is expected to enter into force around August 2016 [3]. Replacement will be permitted for the following mirror classes:

- Class I (interior rear-view mirror) and III (main exterior rear-view mirror) used primarily for passenger cars.
- Class II (main exterior rear-view mirror), Class IV (wide-angle exterior mirror), Class V (close-proximity exterior mirror) and Class VI (front mirror) mainly used for commercial vehicles.

This will establish an international regulatory framework out of a combination of normative and regulatory guidelines. UN Regulation No. 46 has been adopted by 42 countries including substantial vehicle markets such as the European Union states and Russia. For other countries, the use of CMS has to be assessed on a country-by-country basis. Chapter "Standardization and Vehicle Regulation Aspects of Camera Monitor Systems" gives a detailed overview of this situation.

1.2 *Characteristics of Conventional Vehicle Mirrors*

Before beginning with the description of CMS and their potential, this subsection presents the characteristics of conventional mirrors that are of interest for the subsequent argumentation in this chapter emphasizing the passenger cars scenario. Detailed discussion of other vehicles can be found in the special commercial vehicle chapters of this book beginning with chapter "Vision in Commercial Vehicles with Respect to CMS".

The side-view mirror is a mandatory device in today's passenger cars and every driver is familiar with its use. The use of the mirrors is also part of the curriculum in driver's education. A conventional side-view mirror (also called "main exterior rear-view mirror" [1]) is a preventative safety device providing the driver with a view of the road environment behind his ocular points. Figure 3 shows the mandatory minimum required field of view described in UN R.46 for Class III

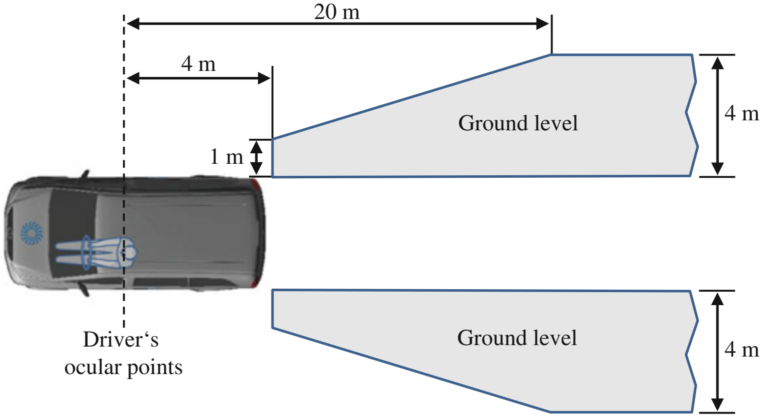


Fig. 3 UN R.46 Class III field of view

devices (the dimensions of the figure are not a true-to-scale representation). The side-view mirrors are specified to provide the driver a greater view of the vehicle's environment than would be possible relying only on the field of view provided by the eyes. Such mirrors must be adjustable from inside the vehicle while the door is closed. UN R.46 uses a vehicle definition to differentiate between categories; according to the definition, a passenger car is an M1 vehicle [1]. The area that is not visible to the driver is called the "blind spot" and depends on the construction of the vehicle.

The specified field of view for Class III (see Fig. 3) can be obstructed under certain conditions by up to 10 %. Obstruction is allowed if this is caused due to the bodywork and its components, such as other cab devices for indirect vision, door handles, outline marker lights, direction indicators and front and rear bumpers, as well as reflective-surface cleaning components [1]. This 10 % permission applies also to the Classes II, IV, V, and VI (used in commercial vehicles). For the Class I mirror, the total obstruction value can be up to 15 % of the mandatory FOV. Obstruction is allowed by devices, e.g., such as sun visors, windscreen wipers or heating elements. For the calculation of the obstruction value, all obstruction devices are considered in sum. According to UN R.46, headrests or framework or bodywork such as window columns of rear split doors and rear window frame shall be excluded from the calculation [1] meaning that in real-life scenarios the obstruction can be much larger. It must be noted that a Class I mirror is mandatory only if the vehicle has a back window large enough to enable seeing the required field of view [1]. The total obstruction of the Class I FOV can also occur in daily use when large persons sit in the vehicle's back seats or if the vehicle carries a large payload. Figure 4 shows the mandatory minimum required FOV described in UN R.46 for Class I devices also called interior rear-view mirror (the dimensions of the figure are not a true-to-scale representation).

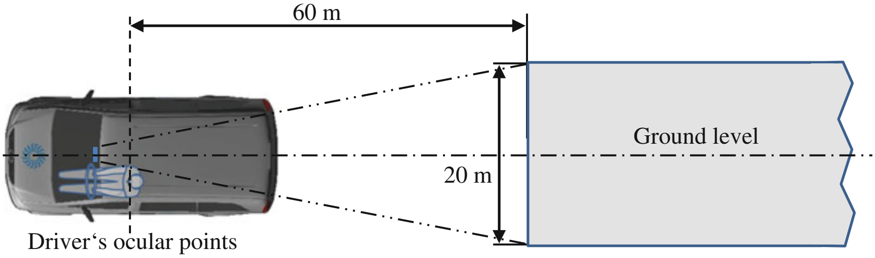


Fig. 4 UN R.46 Class I field of view

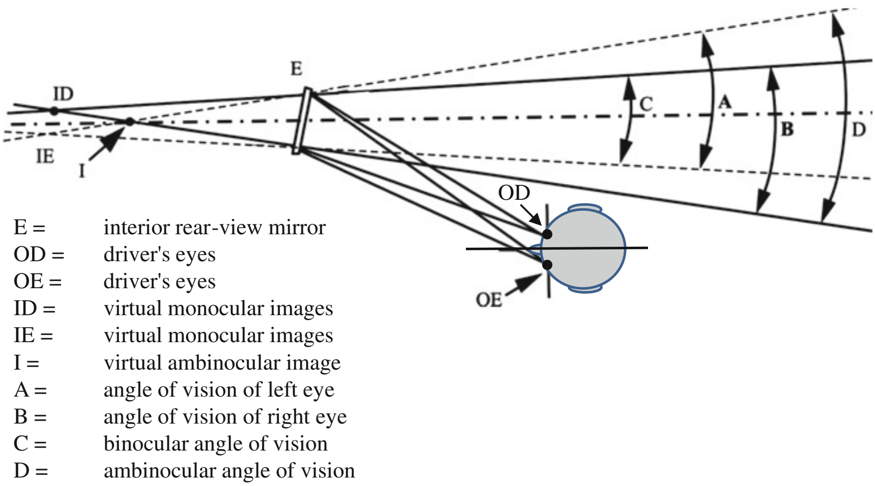


Fig. 5 Ambinocular vision based on UN R.46 [1]

The required fields of view for all Classes are defined using ambinocular vision, see Fig. 5. According to [1], ambinocular vision defines the total field of view obtained by the superimposition of the monocular fields of the right eye and the left eye.

According to UN R.46 the use of spherical convex mirrors is permitted. Such mirrors are defined by the radius of the sphere that specifies the shape of the spherical mirror surface with the reflective layer on the convex surface, leading to positive values for the radius. A spherical mirror has a surface with a constant and equal radius in all directions. For Class I and III, the radius is defined as $r_{mirror/min} = 1.2$ m [1]. For the US market, FMVSS 111 only allows plane mirrors where $r_{mirror/min}$ is infinite on the driver's side. On the passenger's side FMVSS 111 allows a spherical convex mirror with a minimum radius of $r_{mirror/min} = 0.889$ m [2]. In addition, it is possible for the mandatory mirror to extend the field of view by using a aspherical mirror consisting of a spherical and an aspherical portion. In such cases the transition of the reflecting surface from the spherical to the aspherical part has to be marked, and the radius of curvature of the aspherical part shall not be less than 150 mm [1].

The FOV requirements for Class III, as shown in Fig. 3, require a minimum horizontal angle of about 12° be visible to the driver. Typical vehicle mirrors have a larger horizontal angle plus an aspherical part of an additional 10°–15° and in some vehicles can reach a combined angle of up to 35° in the horizontal direction. Through head movement in combination with the movement of the upper body, a driver is able to expand the mandatory field of view for special driving situations such as merging lanes on a freeway.

A driver uses the Class I and III mirrors to simultaneously be aware of the situation behind them while remaining aware of the forward situation seen through the windscreen. Every driver has their own habit for using these devices. Research results have shown that the duration of a single glance at a mirror while performing a forward driving task can be an average 0.3–1 s, depending on the driver and the situation [4]. Additional literature citations concerning ergonomic aspects and the use of mirrors for distance estimation can be found in this book in chapter "Camera-Monitor Systems as a Replacement for Exterior Mirrors in Cars and Trucks".

The described characteristics of a mandatory mirror relate to the intended function to provide indirect vision for a specific FOV. These characteristics can be considered to be the primary function. To support this primary function, some mirrors may also include automatic dimming technologies that are not mandatory. Typically, a side-view mirror can be folded in, either manually or electrically. The implemented electric folding mechanism can be activated by the driver or used automatically depending on, e.g., the vehicle's speed. In many passenger cars, the folding of the Class III mirrors begins when the driver unlocks the car or opens the door to enter the vehicle. For Class I interior mirrors there are tilt options available to manually change between different degrees of reflection for "day" and "night" situations. Such tilt functions are not mandatory according to UN R. 46 and are used to reduce the brightness and glare of lights that is reflected into the driver's eyes. If such tilt functions are implemented, UN R. 46 requires that the "day" position shall allow the colors of the signals used for road traffic to be recognized. The value of the normal coefficient of reflection in the "night" position shall be not less than 4 % [1].

Taking a close look to a mirror of a modern passenger car reveals additional secondary functions of a mirror. After decades of mirror development and installation on vehicles, a mirror can be considered to be a fundamental part of the vehicle. The technological development of the past years has led to a successive use of the mirror housing also as a housing to integrate additional functions and modules. All modern vehicle mirror applications that are not mandatory can be considered to be the secondary functions. The exposed position, e.g., of the side-view mirror housing is ideal for the integration of lamps for illuminating the area next to the door and for the integration of cameras of systems providing a 360° surrounding view of the vehicle. To improve the communication characteristics for wireless services, antennas are also integrated into the mirror-housing, allowing for antenna diversity. Additionally there are heating elements for the mirror surface,

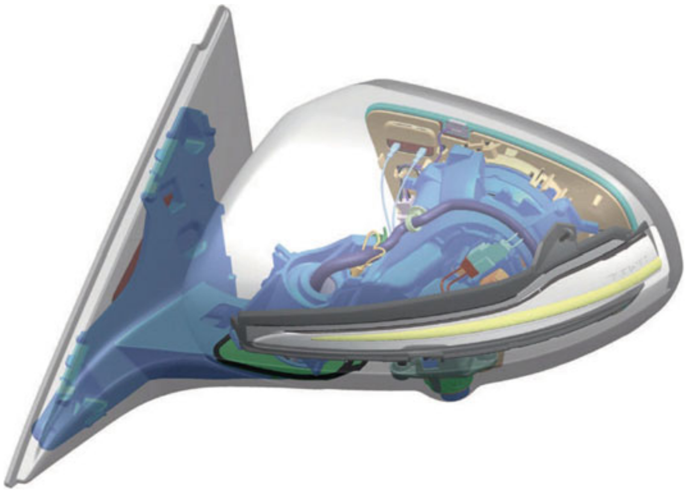


Fig. 6 Side-view mirror of a Mercedes-Benz S Class (2013) showing internal components [5] p. 54 “With permission from Springer” © Springer Automotive Media 2013

dimming electronics, electrical and mechanical components for the folding of the mirror and lamps for the turn signal implemented. Figure 6 shows the side-view mirror of the Mercedes-Benz S Class (2013) displaying the internal components.

Mirrors such as shown in Fig. 6 are aerodynamically and aero-acoustically optimized parts of the vehicle concept. The side-view mirrors are also used to visualize the warning signal of a blind-spot assistance system. When such systems detect a critical object in the blind spot by means, e.g., of radar, a warning signal appears in the glass of the side-view mirror (usually integrated in the non-mandatory aspheric part of the mirror). In some vehicles, the blind spot assistance system is realized using a camera and not radar. The cameras of these systems are also integrated within the side-view mirror housing.

Another secondary function of the mirror is its use as an elementary exterior design element of the vehicle’s body. Vehicle designer would have much greater freedom in designing the vehicle without having the restrictions imposed by having to integrate relatively large mirrors.

Soiling of the side view windows is caused by the air flow around the side mirrors [6]. Current side mirrors are optimized to minimize this soiling. The phenomenon is also dependent on the velocity and cannot be completely eliminated for a side mirror. The following image shows the side-view window at 120 km/h for a vehicle with a very well designed side mirror. The track of dirt and the weakly defined spray above the mirror do not degrade the critical fields of view *a* side-view mirror field of view and *b* eye field of view (see Fig. 7).

UN R.46 also defines requirements concerning the allowed vibration of the mirrors. The attachment onto the vehicle has to guarantee that the mirrors do not move enough to significantly change the field of view, or vibrate to an extent which

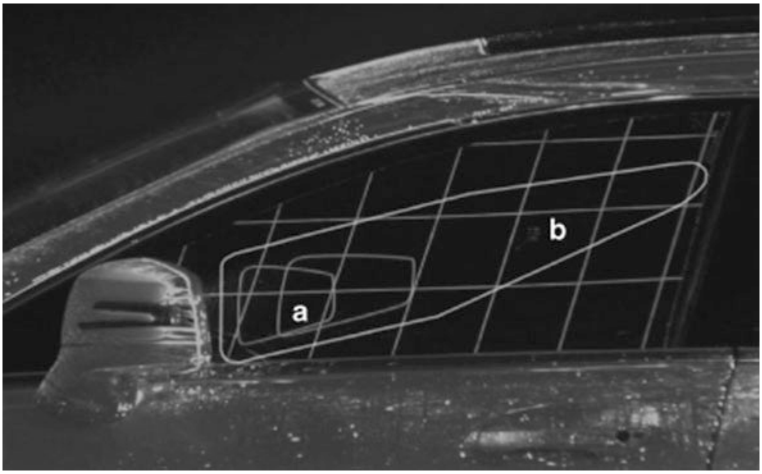


Fig. 7 Side window soiling of a passenger car at 120 km/h [6] p. 471 “With permission from Springer” © Springer Fachmedien Wiesbaden 2013

would cause the driver to misinterpret the nature of the perceived image. This requirement shall be fulfilled when the vehicle is moving at speeds of up to 80 % of its maximum design speed, but not exceeding 150 km/h [1]. To achieve this, depending on the vehicle construction, additional strengthening of the doors is necessary, leading to additional costs and weight. This door strengthening would not be necessary in some vehicle concepts without a mirror at this position.

The complete list with all requirements for all mirror classes can be found in UN R.46 [1]. The requirements are divided into component and installation requirements for the different vehicle categories and also include elements for the protection of pedestrians in the event of an inadvertent contact.

2 Potential for Mirror-Replacement

When analyzing the characteristics of conventional vehicle mirrors as described in Sect. 1.2, there are some points that seem to have obvious potential for improvement, such as a large field of view and other points that seem to be challenging, including the mounting position of the displays in modern passenger cars. The following list summarizes the potential benefits but also the potential challenges of the mirror-replacement by CMS from today’s point of view.

Potential Benefits:

- Improvement of fuel economy;
- Reduction of CO_2 emissions;
- Improvement of aerodynamic and aeroacoustics;

- Improvement of field of view;
- No glare at direct sunlight or true other high-beam headlights;
- Improved night vision;
- Improved vision during rain;
- Improvement of direct vision;
- Reduction of side window soiling;
- Weight reduction of the vehicle;
- Extended range for electric cars;
- Higher maximum speed (for sports cars);
- Helpful to drivers with diminished upper body range of motion;
- Enabling new advanced driver assistance systems;
- More degrees of freedom in the vehicle design;
- Reduction of variants (different mirrors for different countries);

Potential Challenges:

- Integration of the displays within the vehicle (especially for passenger cars);
- Cost level compared to a conventional mirror;
- Modification of the field of view through head movement;
- Near/Far accommodation of the eye (multifocal spectacle glasses);
- Adaptation speed of the CMS at very fast changes of the light conditions;
- Alternative position for the components of today's mirror housing: e.g., surround view cameras, antennas or lamps;
- Ignoring or reducing shoulder check?
- Visibility of the driver in the mirror;

For every different vehicle concept and market situation, all potential benefits and challenges have to be investigated one by one in as much detail as required by the series-development process of the automotive industry. Only a system-level approach can lead to a final conclusion on the relevant points and the priorities associated to each one. Extensive tests of CMS installed in test vehicles (passenger cars and in commercial vehicle) has been carried out by the German Federal Highway Research Institute (BASt). The details and the outcomes of these tests are discussed in chapter "Camera-Monitor Systems as a Replacement for Exterior Mirrors in Cars and Trucks".

2.1 Potential Benefits

2.1.1 Aerodynamic and Fuel Economy Aspects

The automotive industry is currently working on meeting requirements relating to the reduction of CO_2 emissions and the reduction of the fuel consumption of new vehicles. One part of these requirements is the roadmap of the European Commission presented in 2011 with the title "A roadmap for moving to a competitive low carbon economy in 2050" and including milestones up to 2050 [7]. In addition to such roadmaps there are also voluntary commitments made by the automotive industry. A key element of the technical efforts is an emphasis on reducing vehicle weight and reducing air drag. Looking at vehicles as a whole allows combining a variety of measures to optimize overall vehicle design. One of these is the improvement of aerodynamic characteristics by replacing side-view mirrors with CMS. The main effects can be achieved by replacing the Class III side-view mirrors on passenger cars and the Class II and Class IV on commercial vehicles.

The mirror replacement is nothing new from a technical point of view, but until now there was no international legislation framework permitting the use of such solutions for vehicles in series production. For example, a Mercedes-Benz concept vehicle, the F 200, featured a camera monitor system instead of a mirrors as early as 1996-quite a futuristic concept at the time. In more recent years, a number of automotive companies have developed concept vehicles which included camera monitor systems. This is possible because today there are production-ready automotive components (e.g., cameras and displays) for such systems available at an acceptable price level. The Volkswagen XL1 is already in small-scale production and is capable of minimal fuel consumption thanks to a variety of measures that help optimize the overall vehicle design. This concept includes a CMS to replace the Class III side-view mirrors and the authorization followed a special exemption procedure [8]. This method of replacing mirrors can be particularly beneficial for driving cycles such as those typical of larger commercial vehicles. With commercial vehicles, too, mirror replacement systems are just one of a number of measures that can be combined to optimize CO_2 emissions. Detailed discussions concerning CMS for commercial vehicles can be found in this handbook in chapters "Vision in Commercial Vehicles with Respect to Camera Monitor Systems", "Camera Monitor Systems Optimized on Human Cognition-Fundamentals of Optical Perception and Requirements for Mirror Replacements in Commercial Vehicles", "Ergonomic Design of Camera-Monitor Systems in Heavy Commercial Vehicles", "Camera-Monitor Systems as a Replacement for Exterior Mirrors in Cars and Trucks" and "CMS Concept for Commercial Vehicles: Optimized Fuel Efficiency and Increased Safe Mobility".

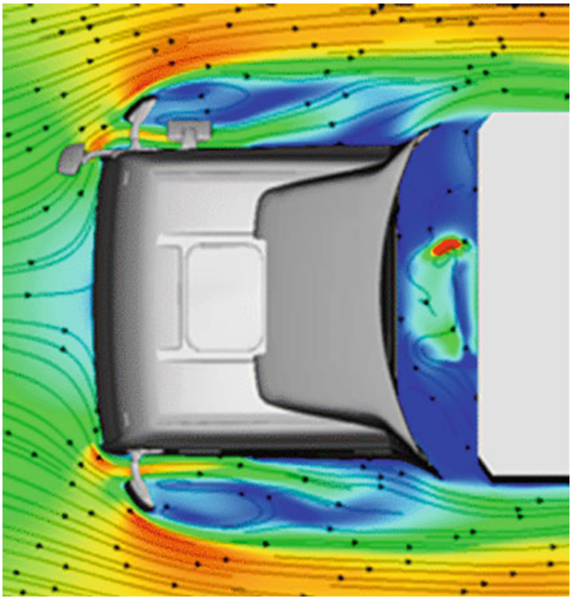
There are a variety of ways to improve vehicles by replacing conventional side-view mirrors with digital mirrors (CMS). Optimizing the aerodynamic properties of the vehicle provides one way to reduce CO_2 emissions and fuel consumption by reducing the air resistance of the vehicle. A detailed discussion of this relationship and effects can be found in [9]. This reduction also depends on the driving cycle and the overall vehicle concept. The air resistance is also called drag and in fluid dynamics refers to the Force F_D acting opposite to the relative motion of the moving vehicle with respect to the surrounding fluid. Depending on the vehicle's shape, about 5 % of the drag of a passenger car is caused by the side-view mirrors [9]. The drag increases as a function of the frontal area A (cross sectional area) of a vehicle and as a function of the vehicle's velocity v. The acting force can be described by the drag equation

$$F_D = \frac{\rho}{2} \cdot v^2 \cdot A \cdot c_W \tag{1}$$

where F_{D} is the drag force, ρ is the density of the fluid and c_W (in literature also called C_D) is the dimensionless drag coefficient [9]. The replacement of the side-view mirror reduces the cross sectional area A of the vehicle and optimizes the drag coefficient c_W. This assumption applies when the camera housing is designed in an appropriate manner that improves the parameter A and c_W in comparison to a conventional side-view mirror. The state of the art implementations of CMS underscores that this assumption is true [8, 10]. The physical dimensions of an available camera are shown in Fig. 28 and are much smaller compared to a side-view mirror, especially compared to the relative large mirrors of commercial vehicles.

Truck mirrors have a peculiarity which needs to be mentioned in this context. Due to their redirecting function, truck mirrors can actually improve the air resistance of the vehicle. Because of the unfavorable flow characteristics of the truck's basic square form, the mirrors can be used in some truck designs to guide the flow along the vehicles side panels [11]. However, this phenomenon cannot be assumed for every truck and only applies for very specific assumptions. Figure 8 graphically illustrates this flow redirection. If the flow meets the vehicle's outer shell behind the A-pillar, though, then the mirrors also increase the c_W value of the truck, depending on the design, by 2–5 % [11].

Fig. 8 Fluid flow through mirror for trucks [11] p. 685 "With permission of Springer" © Springer Fachmedien Wiesbaden 2013



For concept related direct flow of the mirror areas (Class II and Class IV of a commercial vehicle), the theoretical achievable fuel saving potential of a CMS is estimated to be up to 2.9 % [12]. For a series-production commercial vehicle in customer use, the real reduction in fuel consumption through a CMS can be assumed to be 2 %, leading to a reduction of the fuel costs of about 1326 € per year [10].

The drag coefficient of a side mirror itself is relatively large [13]. If this is applied to the complete front surface of the vehicle (for the case of two Class III mirrors), drag coefficients on the scale of c_W 0.008–0.020 result, approximately five percent of the total resistance of a passenger vehicle [13]. If the mirrors are replaced by cameras, an improvement of the c_W value of up to $\Delta c_W = 0.010$ is expected, dependent on the shape of the camera holder [13]. Figure 9 shows the pressure distribution over the front surface of the vehicle. Every vehicle has its own individual pressure distribution because it not only depends on the design of the vehicle front but on complete shape of the vehicle [14]. The relative influence of the side mirrors is also visible in Fig. 9.

Assuming an improvement of $\Delta c_W = 0.010$ for a passenger car, the ECE fuel consumption would be reduced by 0.04 liters/100 km. In everyday use, the reduction can be assumed to be 0.1 l/100 km and for very fast motorway speeds up to 0.4 liters/100 km [15, 16]. To achieve the same level of savings through lightweight-construction measures, a weight reduction of the complete passenger car at least by 35 kg would have to be achieved [16]. In the overall fuel economy calculation, the electric power consumption of all CMS components and all weight elements must be considered. The electric power consumption has negative effects, meaning that an additional power requirement leads to increased fuel consumption and the weight factor has a positive effect. With today's technology, a CMS can be

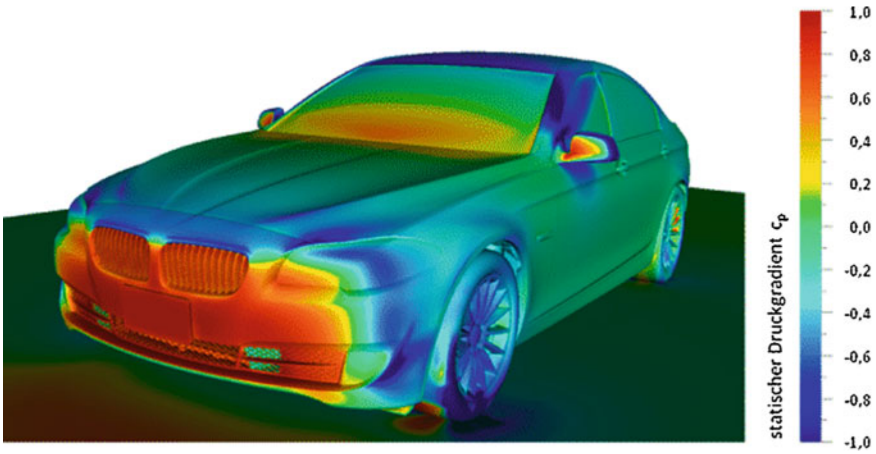


Fig. 9 Pressure distribution over the front of the vehicle [14] p. 510 "With permission of Springer" © Springer Fachmedien Wiesbaden 2013

considered to be much lighter compared to a conventional mirror module. Additional strengthening of the doors needed for the mirrors are not needed for a CMS implementation in some vehicles. The overall effects can be assumed to be positive, leading to better fuel economy with CMS. With e-vehicles, it would be possible to extend the driving distance and with sports cars the maximum achievable speed could theoretically be raised due to the improved aerodynamics.

The aeroacoustics are related to the aerodynamics. A number of aeroacoustic investigations performed during vehicle development are concerned with the side mirrors. They are attached in areas of high air current flows and are therefore extremely critical acoustically. For vans, for example, they can be the primary acoustic source for aerodynamic exterior noise [9]. Especially in the premium segment of passenger cars, a CMS measure can audibly improve the aeroacoustics for passengers.

2.1.2 Potentials for Vision

A key safety aspect which can be improved through CMS is vision (including direct and the indirect vision). For example, the field of vision laid down under UN Regulation No. 46 for a passenger car (Class III side-view mirror) starts on the left and right sides at 4 m behind the eyes of the driver. This can result in invisible zones ("blind spots") that CMS can capture and display by widening the viewing angle using a wide angle camera. In Germany, it is assumed that around 9500 serious road accidents are caused by motorists who are not sufficiently aware of the traffic behind them when changing lanes or cutting off another vehicle too soon after passing. Crash records were extracted for the time period 2002–2006 in order to investigate the crash avoidance potential of five driver assistance systems including blind spot assistance. Figure 10 shows the percentage average numbers of annual crashes within the analyzed time period that are relevant to five driver assistance systems that could potentially have prevented the crash [17].

Blind spot detection including a warning has the potential to prevent 12 % of the crashes. Assuming that CMS can significantly reduce the blind spots or act as a conventional blind spot assistance system (including automatic detection and warning signals) then these crashes could potentially be prevented. This theoretical assumption has to be intensively investigated once CMS is available with a significant market penetration.

The direct view is also limited due to the vehicle chassis, especially due to the pillars. Figure 11 shows an example where the A-pillar, in combination with the side-view mirror, obstructs the direct view to a side-street. In this example, the complete side-street is not visible to the driver, leading to a danger situation since traffic could approach from there.

By replacing the mirrors, the direct view as shown in Fig. 11 can be improved. For this effect it is important to choose an appropriate display position and relatively small camera holder, or to completely avoid the installation of the CMS components within the direct view of the driver.

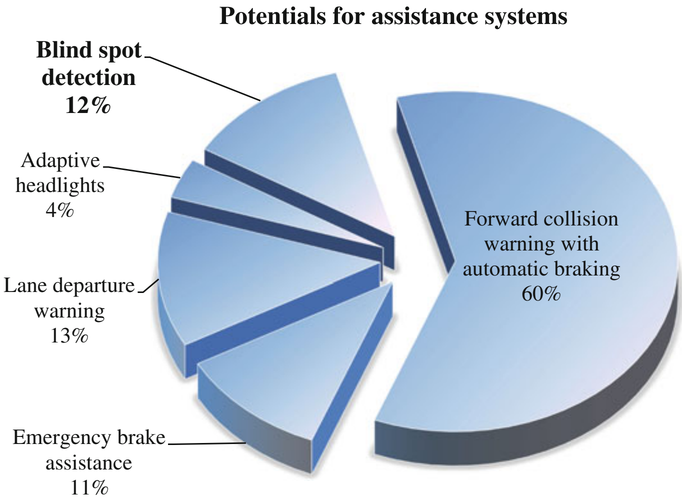


Fig. 10 Potentials for different driver assistance systems showing the blind spot potential based on data from [17]

Fig. 11 Obstruction of the direct view through a mirror, a complete side-street is not visible (*left*)

As shown in Figs. 3 and 4, the UN R.46 fields of view "allow" blind spots next to a passenger car and up to 60 m behind the car. For a truck trailer combination the area covered by the mandatory UN R.46 mirrors are shown in Fig. 12 including the blind spot areas and different positions of the trailer. The blind spots caused by the A-pillar and by the mirrors are clear recognizable (see Fig. 12, white interruption of area 1).

With conventional mirrors, drivers are not able to survey the whole near surrounding area of their vehicle. In the year 2006 it was estimated that every year, over 400 people lose their lives in the European Union because of these kinds of accidents involving trucks. Most victims were pedestrians or two-wheelers, both a particularly vulnerable category of road users [18]. The fields of view for Class V

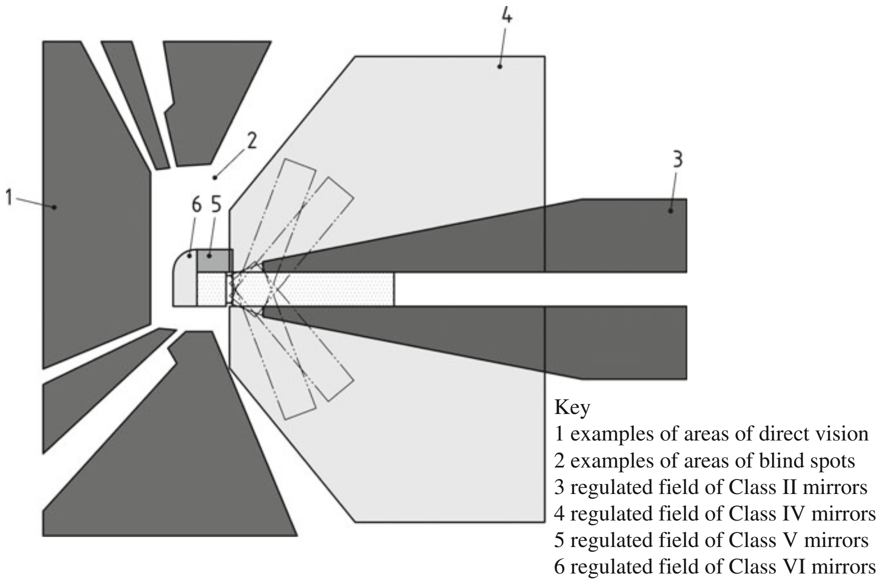


Fig. 12 Area on ground covered by different means of vision for a truck based on [2]

and Class VI of a commercial vehicle addresses exactly these critical situations. But in addition, future systems will be able to warn if objects such as a pedestrian or a bicyclist are next to a truck. Such warning technologies can also be realized by a CMS (Class II and Class IV) in combination with wide angle cameras and appropriate detection algorithms.

The blind spot behind a passenger car has resulted in relatively high numbers of accidents. In the USA alone, about 3000 non-traffic fatal incidents involving children occurred in the time period from 1991 through 2012. About 1100 of these incidents occurred while the vehicle was moving in reverse. The National Highway Traffic Safety Administration published a rule to amend FMVSS 111 in the USA to require all new vehicles (up to 4536 kg) to be equipped with rear-view cameras by May 1, 2018 (beginning on May 1, 2016 with a percentage). An appropriate CMS can increase the field of view of the Class I mirror to also function as a rear-view camera behind the car, e.g., in maneuvering situations. In addition, it is possible to integrate object detection technologies that warn or even brake in critical maneuvering situations.

Reduction of mirror variants is also a potential benefit of CMS for the vehicle manufacturer. Today there are different field of views and mirrors for different countries e.g., for the US market compared to the EU market. CMS allows the adaptation to different countries or markets based on software settings, without the need for hardware changes.

With conventional mirrors, indirect vision can be adapted as required simply through head movement, but this means the driver is constantly moving around.

The need for drivers to move around can be reduced by showing the optimized field of view and combining this with ergonomically positioned displays and display sizes. Especially for drivers with diminished upper body range of motion, such a technology also has the potential to reduce the need for head and upper body movement while driving.

Conventional mirrors also sometimes distract drivers on account of glare from the sun or vehicles behind the car. Glare can be significantly reduced with a CMS. Another benefit is the improved vision during night and rain compared to a conventional mirror (see chapter "Camera-Monitor Systems as a Replacement for Exterior Mirrors in Cars and Trucks" of this book).

The base technology also exists to adapt and optimize views depending on different driving situations. For example, image processing algorithms can be used to recognize dangerous situations and objects such as quickly approaching vehicles and give the driver an early warning and an optimized field of view (including graphical overlays) for this situation. This could make CMS a fundamental technology for the development of other advanced driver assistance systems.

2.2 Potential Challenges

As with every new technology, CMS provides potential advantages and disadvantages. The potential advantages were already discussed in the previous subsection. The potential challenges posed by CMS are discussed in the following.

Currently, the vehicle mirrors are seen as elementary components of a vehicle. Designers consider mirrors during the design of the silhouette of a vehicle as a matter of course. Without side mirrors, designers would have considerably more design possibilities during the design of the vehicle's exterior, which is a clear advantage of CMS. However, there is a challenge regarding the vehicle's interior design. In an initial CMS introduction, customers would expect the CMS displays in the same locations they are accustomed to from the mirrors. This means close to the A-pillar (e.g., in the door panel) in the case of side mirrors as shown, e.g., in Fig. 1. This was verified by a study with test persons (see chapter "Camera-Monitor Systems as a Replacement for Exterior Mirrors in Cars and Trucks"). The development of a vehicle takes approximately six years, and a model upgrade can take place every three years. If current vehicles were converted to CMS, it is possible that suitable space for the display integration may not be available. This could lead to this innovative technology being available relatively late in certain vehicles, depending on the vehicle and development cycle. If the vehicle is offered with conventional mirrors and CMS only as optional equipment, this increases the number of variations for the vehicle manufacturer, although a presumptive advantage of CMS was its potential thereof. There are clear specifications regarding the image size of the displays for a CMS. Thus displays are required which have similar visible areas as mirrors today. If a larger field of view is to be displayed, which is a clear advantage of CMS, then that would require the displays to be correspondingly larger, making the integration even more difficult. As mentioned previously, the field of view for a mirror can be expanded through head movement. This is not possible with a display. If the display therefore showed only the required field of view, this would be a distinct advantage for a mirror. This can be compensated by displaying a larger field of view than the minimum required one, which can lead to larger displays. It is not possible to say with certainty whether the introduction of CMS with a large field of view will eliminate the glance over the shoulder. The changeover to a new system, i.e., from mirrors to CMS, always requires a certain familiarization period. Studies have shown, however, that this familiarization is not critical with regard to driving safety (see chapter "Camera-Monitor Systems as a Replacement for Exterior Mirrors in Cars and Trucks"). The integration of the displays is significantly easier for vehicles designed exclusively for CMS and without side mirrors. In the case of trucks it is easier to integrate the displays compared to a passenger vehicle because there is simply more space in a truck, for example at the A-pillars.

A display is in principle a 2D system (with the exception of special 3D displays). Thus it cannot be used to measure stereoscopic depth information. In return, it must be considered that the possible stereoscopic depth perception due to the curvature of the mirror may be lost. Further, stereoscopic depth perception for humans is only relevant for a relatively short distance of a few meters. For viewing distances such as those which result for the field of views of mirrors, depth perception takes place using other visual possibilities which are most definitely also included in the display image. That explains why distances as well as speeds can be judged comparatively well with a CMS as with a conventional mirror. A detailed analysis of the visual perception of humans is included in chapter "Human Visual Perception".

Drivers who wear multifocal lenses must also be considered when determining the position of the displays. For these drivers, it is advantageous to have the displays mounted somewhat lower in order to facilitate the readability. ISO 16505:2015 allows a maximum downward angle of 30° for this purpose. The human eye must focus relatively quickly on a short distance, e.g., 60 cm when glancing at the display. The situation in front of the vehicle is seen in focus several meters ahead of the vehicle. This transition is called the near/far accommodation of the eye. The ability of the eye to make such transitions quickly declines with age, and this must also be taken into consideration when introducing CMS. There will most certainly be drivers who, due to their personal aptitude and inclinations, will prefer to continue using a conventional mirror.

A further aspect to bear in mind in the comparison with conventional mirrors is the price of a CMS. A mirror is principally a passive system. A CMS consists of cameras, displays, and a corresponding control unit with processing capabilities. Although these components have reached a price level that allows their use in series systems such as driver assistance systems, in a direct comparison they are still generally more expensive than a mirror. As already discussed in Sect. 1.2, additional components such as, e.g., cameras, antennas, or peripheral lighting are already being integrated into mirror housings (see also Fig. 6) and these would require integration elsewhere if side mirrors were eliminated.

The driver is visible in a conventional side mirror. Especially for large trucks, this can serve as a type of communication channel between the driver and bicyclists. This possibility does not exist in the case of CMS.

Besides the purely objective potential advantages of a CMS, there is of course the possibility that an emotional vehicle design, made possible through CMS, can be an influencing factor in the customer's decision to buy even if the vehicle is more expensive as a result. This phenomenon should not be underestimated, especially in the case of luxury vehicles. CMS is a visible and perceptible innovation.

This book claims to discuss this topic as broadly as possible and also critically in order to present the reader with as complete a picture as possible. Despite all the technological challenges, it is important to emphasize that a CMS can be implemented as a replacement for mirrors and that the discussed advantages can outweigh the disadvantages. First studies have also shown that drivers can drive with CMS with comparable safety to current mirrors (see the study in chapter "Camera-Monitor Systems as a Replacement for Exterior Mirrors in Cars and Trucks").

3 Overview of Camera Monitor Systems

The following section provides an overview of the technical implementation and architectures of CMS. The function of CMS is to generate an optical image signal of the prescribed field of view and to provide it optically to the driver in real time. The driver should see an accustomed image compared to the known conventional mirror.

3.1 Base Architecture of a CMS

As a rule, three main components are required for the basic architecture of a CMS. The image is generally captured by means of a camera. The camera signal is then transmitted to a control unit (ECU) for further processing. The control unit prepares the image so that it is presented optimally in the display. There is a video interface between the control unit and the display as well as an additional interface for transmission of control and monitoring data. There is a comparable interface between the ECU and the camera. Figure 13 shows the block diagram of a CMS.

The displayed architecture can be realized on the basis of three different components which each need a power supply. The control unit could, however, also be integrated into the display housing, leading to an architecture with two components. It is also possible to design the individual processing tasks of the control unit as hardware or as software, or as a combination of HW/SW. In the current automotive environment, the camera is in most cases a module consisting of an optical lens, a digital imager and a so-called serializer, the interface-component. The optics bundle

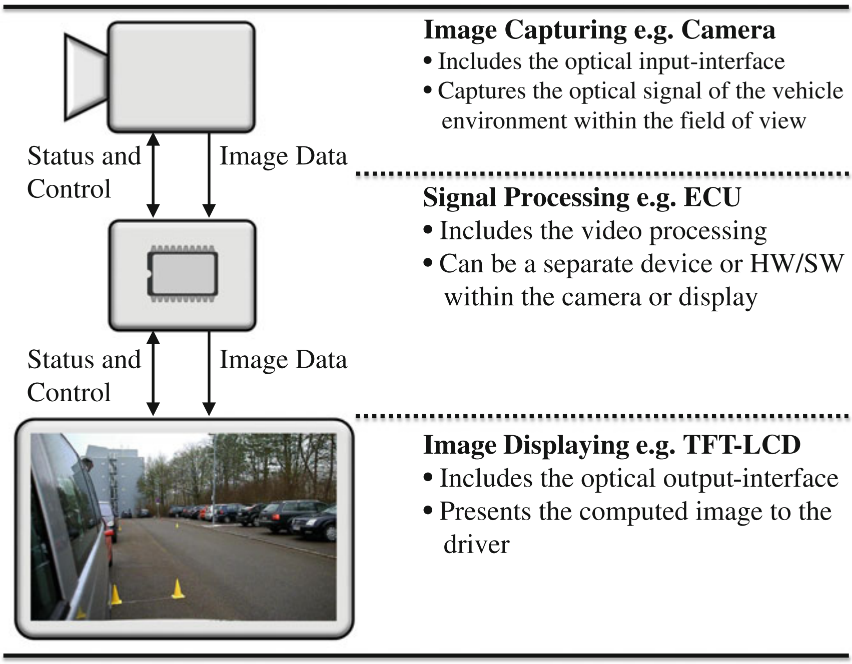


Fig. 13 Camera monitor system block diagram

the incident light rays and project these onto the light-sensitive surface of the imager. Imagers are matrix arrays made up of photodetectors which each evaluate a single image point (pixel), i.e., they deliver an electrical signal proportional to the radiant power incident on that pixel. In addition to the actual electro-optical conversion, the readout procedure in which the image information in the matrix is read out in series, is also of great importance. Currently, imagers with a resolution of 1 megapixel are common in automotive applications (e.g., the resolution is then 1280×800 pixels). Such imagers are now being fabricated using Complementary Metal Oxide Semiconductor (CMOS) technology. The design and basic functionality of such a camera is presented in Fig. 14.

In order to also receive a color signal, each individual photodetector has a color filter inserted ahead of it. As shown in Fig. 14, it is a color filter for green, red and blue. Such filters are called Bayer or Mosaic filters. Thus the imager performs a spectral separation of the incident radiation. The electro-optic conversion takes place on the basis of the photoelectric effect, e.g., using a photodiode. During the exposure time a number of photons hit the pixel area. Due to the photoelectric effect, a number of electrons result that affect a voltage level. For the control of the exposure time, additional electronic circuits are implemented. The voltage is amplified and then digitized by an analog to digital converter. A digital value for each pixel results. There is a relationship between the exposure time, the maximum

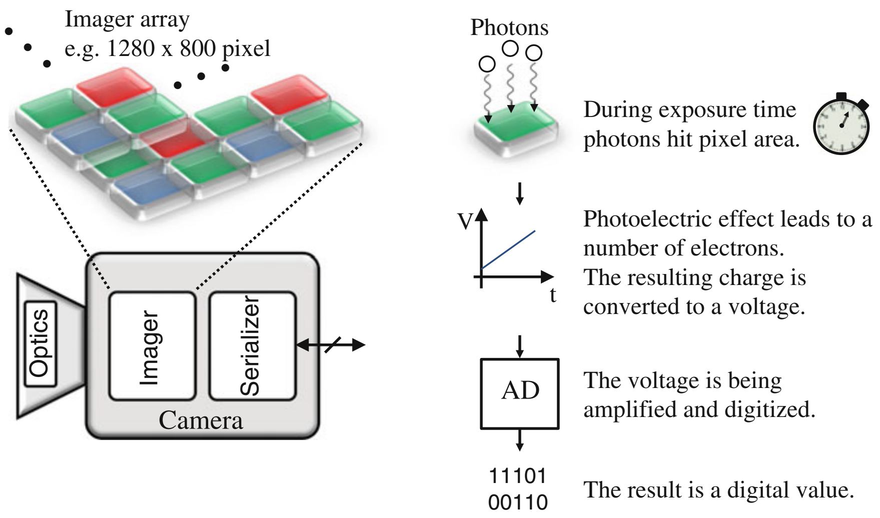


Fig. 14 Camera principle

frame rate and the achievable dynamic range of the camera. Using, e.g., a relatively long exposure time in dark situations (such as at night) leads to a lower achievable frame rate. Parameters like the frame rate of the camera can be controlled using the control interface of the camera. There are different techniques available to use different exposure times in order to achieve a high dynamic range (HDR), which is very important for a camera monitor system. Additional information about the HDR topic can be found in chapters "Image Sensors for Camera Monitor Systems" and "Optimization of Demanding Scenarios in CMS and Image Quality Criteria". Chapter "Image Sensors for Camera Monitor Systems" is dedicated to relevant imager technology.

The optical lens of a CMS can be characterized by its focal length (determines the angle of view), by the aperture, and by the spectral properties. With respect to CMS, there are optical effects and parameters that need to be considered since they can influence the system performance. Such effects include different types of flares, e.g., veiling Glare, directed flares, aperture ghost and ghost images. A detailed discussion can be found in chapter "Optical Effects in Camera Monitor Systems".

A serializer device is used for the data transmission from the camera to the ECU. The imager output is usually a parallel output having 10 or 12 signal lines. For automotive video links, the transmission is not realized by using 10 or 12 separate parallel lines, but by an appropriate serial data transmission using, e.g., shielded twisted-pair lines. A parallel to serial converter is implemented in the serializer device for this task and the serial transmission is realized by low-voltage differential signaling (LVDS). Depending on the frame rate, the video format, and the resolution, the resulting serial video data stream can have a bit rate of around 1 Gbps including a back channel to control the imager using a lower bit rate. Currently there are different interface technologies available including variants in which the camera signal has to be compressed. Additional information about the different video interfaces can be found in chapter "Video Interface Technology".

Within the ECU, there usually is a deserializer device performing a serial to parallel conversion of the data stream. One signal processing task within the ECU is to calculate the color information based on neighboring pixel information, called the demosaic or debayer process, and to adapt the video signal for the display. A one-by-one representation of every camera pixel to every display pixel is recommended. The display can have a different resolution than the camera. In this case, a windowing can be used to display the relevant portion, e.g., containing the required field of view. The display of a CMS can be considered a vital component since it is the direct interface between the driver and the technology.

The embedded technology that can be used for a CMS, including special semiconductor devices and design tools, is described in chapter "Camera-Monitor-Systems as Solution for the Automotive Market.

3.2 Advanced Architecture of a CMS

While the basic architecture is suited for the replacement of individual mirrors, advanced architectures are realizable which contain additional interfaces or which can replace multiple mirror classes in a common system architecture. In the basic architecture, the camera image is only transmitted to the display. In an enhanced architecture, the image signal can also be transmitted to other locations in the vehicle. This is necessary, for example, if the information is processed further by another control unit in a driver assistance system. During the subsequent processing, object-recognition algorithms can be used to identify critical objects such as other vehicles or pedestrians. It is also possible to fuse the image data with data from other sensors so that a blind spot assistant application can fuse camera data with radar data in order to improve the detection quality. Recognized objects can also be transmitted back to a CMS in order to be displayed for the driver. For example, this allows graphical overlays to be superimposed onto the CMS image to present the driver with additional information or warnings. Overlays can also be guidelines which are required for maneuvering and for estimating the distances to objects in the vehicle's surroundings. An example of this architecture, with the respective interfaces, is presented in Fig. 15 and also appropriately described in the ISO 16505:2015 standard.

A control interface can be used to control the complete CMS, e.g., in order to set the zoom factor or to select a specific field of view. If a CMS is integrated into a modern vehicle, interactions with other vehicle systems are an unavoidable consequence. A CMS thus requires interfaces, for example, to the following vehicle systems: gear selection, door control unit, seat control unit, rain sensor, light sensor, and turn signals. These interfaces are typically realized via the vehicle busses

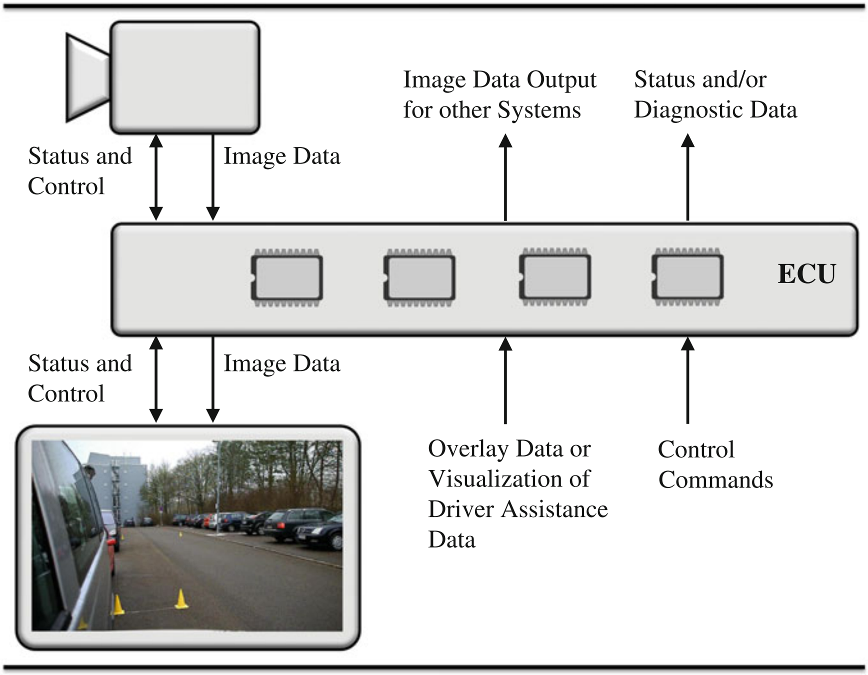


Fig. 15 Camera monitor system block diagram with advanced interfaces

(e.g., CAN or FlexRay). The ISO 16505:2015 standard, as already mentioned, is described independently of a concrete technical implementation. Thus it is possible, for example, to use multiple cameras to capture images. Each camera image can be displayed in a separate display or, as in the case of Class II and Class IV, a correspondingly larger display can be used to present both images simultaneously. In this case, it is necessary to ensure that the images are clearly separated for the driver, e.g., by a separating line in the display.

Since a camera monitor system involves electronic parts which impact vehicle safety, ISO 16505:2015 requires dealing with the system in accordance with safety processes outlined, e.g., under ISO 26262. Especially for enhanced CMS architectures, special attention must be paid to the topic of safety because the degree of integration and complexity are increased significantly. This special topic is the subject of its own chapter in this book (chapter "Functional Safety of Camera-Monitor Systems").

Figure 16 shows examples of a CMS with object recognition and visual output with assisting overlays. The image on the left shows a scene in which a bicyclist is riding to the right of a truck. A Class II camera in combination with recognition and tracking algorithms recognizes the bicyclist and warns the driver through the visual projection of a transparent overlay. In addition, an acoustic warning can be emitted if the situation becomes more critical. The images on the right in Fig. 16 show an

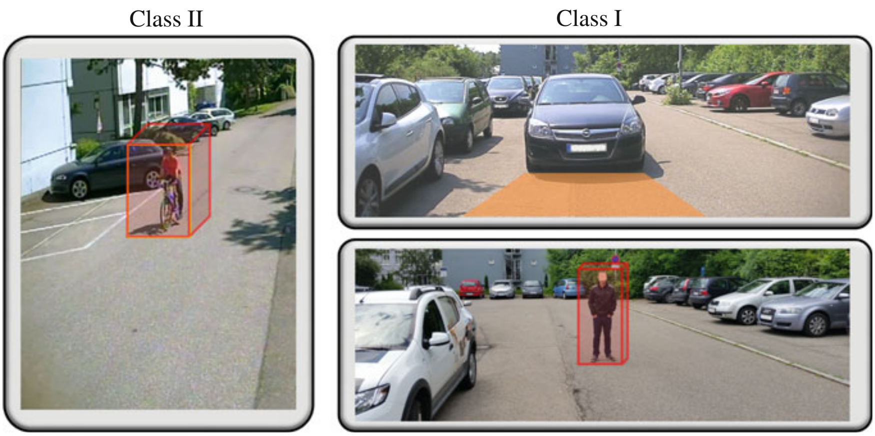


Fig. 16 CMS examples including object detection and warning overlays

example of a Class I rear-view camera. In this case, a person can be detected and tracked. The detected person is marked by a transparent red overlay for the driver. In this case it is also possible to emit an acoustic warning if the situation becomes more critical. The upper Class I example shows a transparent overlay projection in which the color of the overlay can be varied dependent on the distance in order to intuitively assist the driver.

In accordance with ISO 16505:2015, overlays are generally allowed [2]. Within the required field of vision only temporary and transparent overlays are allowed, for example, see Fig. 16 illustrating a detected person. In addition ISO 16505:2015 permits any driving-related visual information to be added to the original image as an overlay without a limitation of the size. UN Regulation No. 46 is stricter with regard to overlays. Any overlay is considered to be an obstruction, including transparent overlays. It also permits only temporary overlays but refers solely to rearward-driving related visual information within the minimum required field of vision [3]. The maximum size of every overlay is 2.5 % of the surface displaying the field of vision. For the calculation of the allowed obstruction of the field of vision, the overlays also need to be considered. UN R.46 permits an obstruction of a maximum of 15 % for Class I and up to 10 % for the other classes.

4 CMS Requirements

Every system designer will define specific requirements dependent on the aspects of vehicle integration and of the overall requirements for the different functions of the CMS. In addition, further requirements are based on conclusions relating to ergonomics which resulted from investigations with test persons. The following subsections summarize the CMS requirements which are based on the relevant standards and regulations as well OEM-specific stipulations. OEM (Original Equipment Manufacturer) in this context refers to the vehicle manufacturers. There may be additional legal and customary market requirements in addition to those mentioned in Sect. 4. Further, the latest developments from science and technology should be taken into consideration for each new CMS project.

4.1 Relevant Standards and Regulations

The ISO 16505:2015 [2] standard as well as United Nations Regulation 46 [1, 3] form the basis for the CMS requirements. Both documents contain cross references to additional standards and regulations which must be met and followed. In an actual CMS project, the OEM will introduce further requirements of their own. Figure 17 shows an example of the standards and regulations as well as the OEM requirements involved in a CMS project. In the OEM requirements, topics such as, for example, EMC (electromagnetic compatibility) may be subject to stricter demands than provided by relevant UN Regulation No. 10.

In addition to the standards shown for ISO 16505 in Fig. 17, the ISO document cites additional standards (e.g., for definitions of terms) which are not explicitly

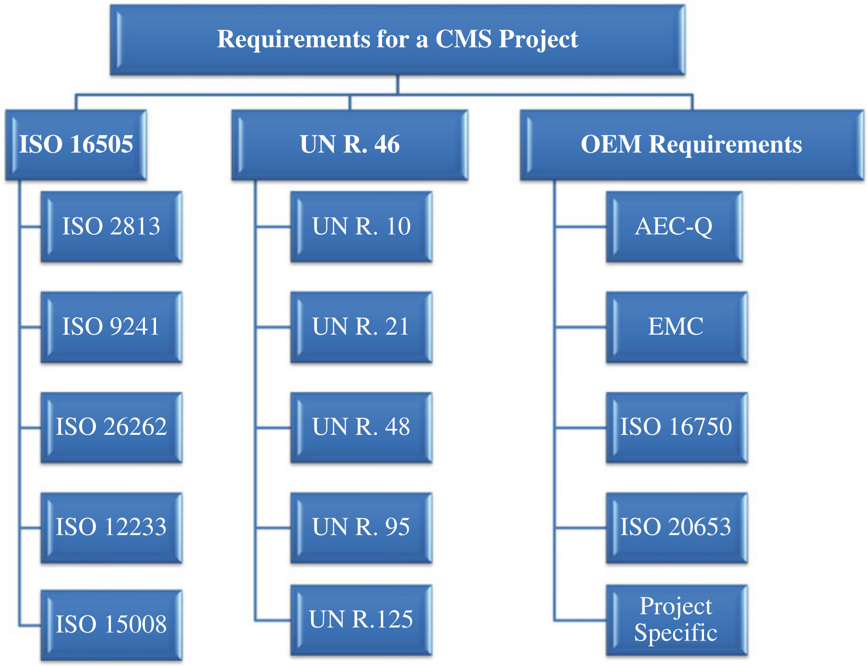


Fig. 17 Example requirements for a CMS project

Table 1 Standards relevant for ISO 16505:2015

Standard	CMS relevance
ISO 2813:2014 paints and varnishes—determination of gloss value at 20°, 60° and 85°	Gloss of the monitor housing
ISO 9241-302:2008 ergonomics of human-system interaction—part 302: terminology for electronic visual displays	Image formation time, monitor coordinate system, luminance contrast
ISO 9241-305:2008 ergonomics of human-system interaction—part 305: optical laboratory test methods for electronic visual displays	Determination of the reflected light, luminous density measurement
ISO 9241-307:2008 ergonomics of human-system interaction—part 307: analysis and compliance test methods for electronic visual displays	Definition of "reality information", visual artifacts
ISO 26262-1:2011 road vehicles—functional safety	Application of a safety standard to the complete CMS
ISO 12233:2014 photography—electronic still picture imaging—resolution and spatial frequency responses	Definition and measurement of spatial frequency response, visual resolution, definition of MTF parameter
ISO 15008:2009 road vehicles—ergonomic aspects of transport information and control systems—specifications and test procedures for in-vehicle visual presentation	Monitor integration, definitions for luminance and contrast rendering, image quality measurement

Table 2 Regulations relevant for a CMS project

United Nations Regulation	CMS relevance
Uniform provisions concerning the approval of vehicles with regard to	
No. 10: "Electromagnetic compatibility"	EMC compatibility of the complete CMS
No. 21: "Interior fittings"	CMS components behavior in case of impact with occupants or pedestrians
No. 48: "Installation of lighting and light-signaling devices"	Position of the direction indicators (e.g., currently part of the mirror housing)
No. 95: "Protection of the occupants in the event of a lateral collision"	Relevant for the characteristics of the in-vehicle CMS components, e.g., displays
No. 125: "Forward field of vision of the motor vehicle driver"	Position of the CMS components within the driver's forward field of view

listed. Table 1 contains an overview of the standards relevant to ISO 16505:2015 with reference to the corresponding requirements.

Table 2 shows the regulations which are relevant for a CMS project in addition to UN R.46.

In the field of automotive electronics in particular, there are external influences and requirements which are not specific to a CMS and are relevant for all electrical and electronic components in the vehicle. For this reason, these requirements were not explicitly included in the ISO 16505:2015 requirements. In order to accommodate environmental requirements, possible failure mechanisms and impacts from environmental influences must be understood. One of the most relevant environmental influences is high temperature, which can lead to diffusion, the formation of intermetallic phases, and to degradation processes. A further, very significant influencing factor is temperature variation, which can lead to material fatigue, delamination and fracturing. Chemical effects lead to ion contamination and resulting corrosion and delamination. Additional factors are physical impacts and mechanical stress. These also lead to material fatigue, fracturing, and delamination. An electric field can result in dielectric breakdown, sticking and latch-up. The evaluation of components for implementation in vehicles through extensive testing is called qualification. The qualification testing can be performed by suppliers or OEMs in accordance with company-specific standards. Uniform processes within industry sectors are advisable. Since these effects are relevant for all automotive electronic components, there are special requirements and standardized test procedures which have been formulated by the Automotive Electronics Council (AEC). Figure 18 provides an overview of the environmental influences which can either simultaneously or sequentially impact a CMS.

The AEC publishes three series of documents for this purpose, the Q100 series for integrated circuits (for example, the imager of a CMS camera), the Q101 series for discrete semiconductors, and the Q200 series for passive components. Thus, for

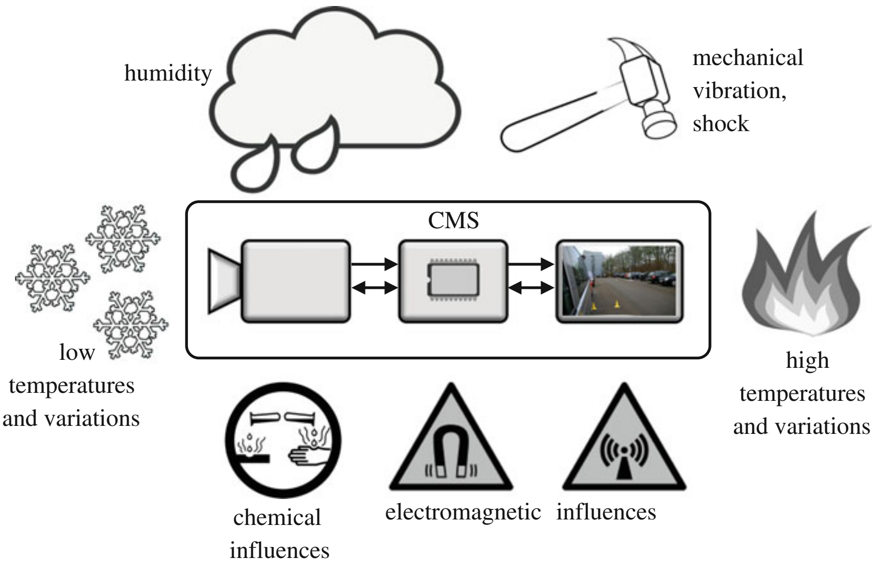


Fig. 18 Environmental conditions for a CMS

example, the following are provided for resistors: temperature and temperature variation tests, moisture tests for operation and storage, lifetime tests, visual inspection, mechanical load tests for the contacts, soldering tests, chemical tests, vibration and impact tests, electrostatic discharge tests, and flammability tests. The temperature requirements for automotive electronics such as the CMS components are typically an operating temperature range of −40 °C to +85 °C and a storage temperature range that can be even higher.

Table 3 contains the OEM requirements which might be relevant for a CMS project. The list shown here contains possible requirements. These can be different for each OEM and might contain either fewer or more requirements. The project-specific OEM requirements may contain further technical parameters, standards, regulations, supplier requirements and specifications for the development process which have not been mentioned here.

Table 3 OEM-requirements relevant for a CMS project

OEM-requirements	CMS relevance
AEC—Q100: Failure mechanism based stress test qualification for integrated circuits	Relevant for the qualification of the electronic integrated circuits of a CMS (e.g., the imager in the camera)
AEC—Q101: Failure mechanism based stress test qualification for discrete semiconductors	Relevant for the qualification of the electronic discrete semiconductors of a CMS
AEC—Q200: Stress test qualification for passive components	Relevant for the qualification of the passive electronic components of a CMS
Electromagnetic compatibility EMC requirements	In addition to UN R.10 OEM specific EMC parameter and test procedures for the CMS
ISO 16750-3:2012 Road vehicles—environmental conditions and testing for electrical and electronic equipment—part 3: mechanical loads	Vibration, free fall, and mechanical shock behavior of the CMS components
ISO 16750-4:2010 Road vehicles—environmental conditions and testing for electrical and electronic equipment—part 4: climatic loads	Operating temperature range, the storage temperature range and behavior in a humid environment under temperature changes of the CMS components
ISO 16750-5:2010 Road vehicles—environmental conditions and testing for electrical and electronic equipment—part 5: chemical loads	CMS components resistant to the different chemical substances. Camera resistance, e.g., to salt spray
ISO 20653:2013 Road vehicles—degrees of protection (IP code)—protection of electrical equipment against foreign objects, water and access	Camera dust-tight and resistant to high-pressure or steam-jet cleaning
Project specific	All additional parameters and aspects that are relevant for OEMs for a CMS project

4.2 Specific CMS Requirements Based on ISO 16505 and UN R.46

ISO standard 16505 defines the minimum ergonomic and performance aspects of camera monitor systems and includes requirements and test procedures for complete CMS. The minimum fields of view to be displayed corresponds, e.g., to UN Regulation No. 46 requirements. One of the system advantages of a mirror compared to a CMS is that everything is in actual real time. Processing tasks involving the camera, data transfer, ECU signal processing and the display can result in slight delays. Under ISO 16505, the system's complete latency time (glass to glass) must be under 200 ms with a video frame rate of at least 30 Hz (or at least 15 Hz in darkness, e.g., in night situations). Another difference compared to a mirror is that a camera and a display present an image with discrete picture elements, known as pixels, and the number of pixels is not infinite. For the resolution requirements of a CMS, ISO 16505:2015 follows a method that considers the visual acuity requirements of the driver including parameters for the camera and display position. Based on the eyes' visual acuity V_{eye} a driver is limited in resolving small details or distinguishing a number of line pairs in a mirror, or on the display in case of the CMS. Obtaining a driver's license requires a minimum visual acuity $V_{eye/min}$. In Germany, for example, the minimum visual acuity is 0.7 to obtain a driver's license for passenger cars and for commercial vehicles like trucks (driving licenses for C/CE) or taxis it is 0.8 [2]. Not every national body requires the same value, e.g., the Netherlands and Sweden stipulate 0.5 for passenger cars. The visual acuity is measured using, e.g., a Landolt C test at an optician and is measured in 1/arcmin. This Landolt C test is described in ISO 8596:2009 "Ophthalmic optics—Visual acuity testing—Standard optotype and its presentation". Glasses or contact lenses are allowed during the test to achieve the required acuity. In this case, the driver's license includes a reference to the glasses. Figure 19 shows such a chart with Landolt C rings (left) and the test of the visual acuity of the human eye (right). The distance between the eye and the test chart is *b* and the Landolt C ring openings are sized so that the openings are visible under an angle of 1 arcmin. If the driver correctly identifies the direction of 60 % of the Landolt C ring openings a visual acuity of 1 results [2].

In terms of image resolution of a CMS, ISO 16505:2015 outlines a calculation method which accounts for the camera and display position as well as for the visual acuity of the driver. A modulation transfer function (MTF) is used as a quality and measurement criteria for image resolution. It defines how many black and white lines have to be resolvable by the complete CMS. The unit of this requirement is line widths per picture height [LW/PH] and has to be calculated for every CMS individually according to ISO 16505:2015. MTF10 defines the spatial frequency of the MTF, where the modulation has dropped to 10 % of the modulation of its reference black and white signal level and MTF50 defines the corresponding frequency at 50 %. The assumption of a monitor with 1:1 aspect ratio is indicated by adding 1:1 to the abbreviation and the MTF values are considered in vertical and

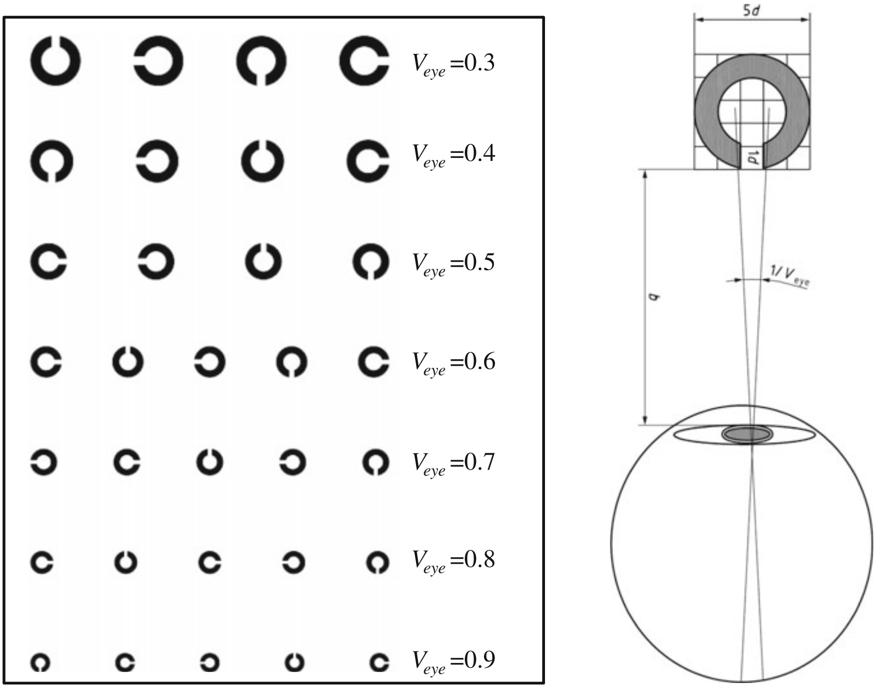


Fig. 19 Chart with Landolt C rings (*left*) and the test of the visual acuity of the human eye (*right*, [2])

horizontal direction within the portion of the monitor that displays the required field of view. A detailed description of this topic can be found in ISO 16505:2015 and in this book in chapter "Resolution and Sharpness Requirements for CMS".

Another system-relevant parameter is the size of the objects visualized in the display. As explained in Sect 1.2, a conventional spherical mirror has a surface defined by a radius. Objects in such mirrors appear smaller than they really are. To consider this situation for a CMS specification, a parameter called magnification factor of the system M_{system} is used that defines the relationship between the correct size of an object and its perceived size when viewed in the mirror [2]. ISO 16505:2015 defines different magnification factors for the different classes based on real values from current homologated vehicles. Following this design method, a CMS will visualize objects in the same way in terms of size as a conventional mirror in current homologated vehicles. Table 4 shows the relevant magnification factors for a Class III passenger car mirror and for Class II and IV commercial vehicles' mirrors.

It is not the intension of this book to cover all parameters described in ISO 16505:2015 or in UN R.46. Selected parameters that are relevant for a system-level understanding of the design method of a CMS are presented. Table 5 shows a summary of relevant parameter of a CMS design.

Table 4 Magnification factors according to ISO 16505:2015 [2]

Magnification factor	Value
Mirror average and minimum magnification factor for Class III UN R.46 mirrors	$M_{mirror/driver/avg}$ = 0.31 (driver side)
	$M_{mirror/driver/min}$ = 0.29 (driver side)
	$M_{mirror/passenger/avg}$ = 0.20 (passenger side)
	$M_{mirror/passenger/min}$ = 0.19 (passenger side)
Mirror average and minimum magnification factor for Class II UN R.46 mirrors	$M_{mirror/driver/avg}$ = 0.23 (driver side)
	$M_{mirror/driver/min}$ = 0.21 (driver side)
	$M_{mirror/passenger/avg}$ = 0.15 (passenger side)
	$M_{mirror/passenger/min}$ = 0.13 (passenger side)
Mirror average and minimum magnification factor for Class IV UN R.46 mirrors	$M_{mirror/driver/avg}$ = 0.065 (driver side)
	$M_{mirror/driver/min}$ = 0.037 (driver side)
	$M_{mirror/passenger/avg}$ = 0.036 (passenger side)
	$M_{mirror/passenger/min}$ = 0.014 (passenger side)

As already mentioned, Table 5 is merely a summary of selected CMS parameters. The complete list and the definitions can be found in ISO 16505:2015 and also includes the test procedures for all mandatory requirements.

ISO 16505:2015 was the basis for the amendment to UN Regulation No. 46 [3]. For some parameters, modifications compared to ISO 16505:2015 were introduced. One example relates to the arrangement of the monitor inside the vehicle. According to ISO 16505:2015, one single central display can be used for Class III and I in a passenger car. This is not permitted by UN R.46 [3] and the possible monitor arrangement is defined to correspond to the mirror positions of conventional mirrors. In addition to ISO 16505:2015, the following list summarizes examples where modifications and additional requirements apply in UN R.46:

- **Point light sources**: The CMS shall have an operation mode in which it is possible to recognize two point light sources (e.g., passing beam headlights in night situations) rendered as two distinguishable separate point light sources [3]. A corresponding test procedure and measurement values are also included.
- **Definition for dual function systems**: A combination of a CMS and a Class I mirror. In such systems, a monitor complying with UN R.46 is placed behind a semi-transparent mirror complying with UN R.46. It is possible for the driver to

Table 5 Selected CMS parameters according to ISO 16505:2015 [2]

CMS Parameter	Requirement/Value
Field of View (FOV)	Display the required FOV, e.g., according to UN R.46
System latency (glass to glass)	<200 ms (at room temperature 22 °C ± 5 °C)
Image formation time (display)	<55 ms (at room temperature 22 °C ± 5 °C)
Frame rate	≥30 Hz (≥15 Hz, e.g., in night situations)
Operating readiness (system availability)	(a) Switch-on-time for a cold start $t_{ON} \leq 7$ s
	(b) $t_{RESTART} \leq 1$ s for a stand-by modus
Average magnification factor in horizontal and vertical direction (the same method also applies to the minimum factors)	$M_{system/hor/avg} \geq M_{mirror/driver/avg}$
	$M_{system/ver/avg} \geq M_{mirror/driver/avg}$
	$M_{system/hor/avg} \geq M_{mirror/passenger/avg}$
	$M_{system/ver/avg} \geq M_{mirror/passenger/avg}$
Resolution in horizontal direction	$MTF10_{(1:1)/hor} \geq MTF10_{MIN(1:1)/hor}$
Resolution in vertical direction	$MTF10_{(1:1)/ver} \geq MTF10_{MIN(1:1)/ver}$
Sharpness in horizontal direction	$MTF50_{(1:1)/hor} \geq ½\ (MTF10_{MIN(1:1)/hor})$
Sharpness in vertical direction	$MTF50_{(1:1)/ver} \geq ½\ (MTF10_{MIN(1:1)/ver})$
Luminance and contrast rendering (values for minimum luminance contrast on the monitor)	(a) for direct sunlight condition: 2:1
	(b) for day condition with diffuse ambient light: 3:1
	(c) for sunset condition: 2:1
	(d) for night condition: 5:1

switch between the two modes (CMS and mirror) and the monitor is visible in the CMS mode.

- **Luminance and contrast rendering**: For night conditions, the value is increased to be 10:1 except in the case of dual function system of Class I where the value is 5:1 [3].
- **Luminance diffuse illuminator**: For the test procedure for daylight conditions with diffuse sky-light exposure, the value of the illuminator is increased from up to 1500 cd/m^2 (in ISO 16505:2015) to up to 4200 cd/m^2 (in UN R.46).

There are also modifications concerning the operating readiness (system availability) of the CMS and the need to perform dedicated EMC tests according to UN R.10. A discussion of the differences between ISO 16505:2015 and UN R.46 can be found in this book in chapter "Standardization and Vehicle Regulation Aspects of Camera Monitor Systems".

5 CMS System Design Based on ISO 16505

In the following section, relevant parameters and considerations are presented for a CMS system design following the methods as defined in ISO 16505:2015. For the system design, there are vehicle and mirror related parameters to be considered. Important CMS parameters are the field of view and resolution of the camera, size and resolution of the display, and the positions of the camera and the display.

5.1 General Considerations and Definitions

A CMS system design starts by considering the driver's ocular points within a given vehicle. The definition of the driver's ocular points can be found in [1]. A so-called *R* point defines a seating reference position for the driver and is defined based on construction data provided by the vehicle manufacturer with respect to a three-dimensional reference system. The driver's ocular reference point (ORP) defines the middle point between the two ocular points of the driver as can be seen in Fig. 20. The reference system of the coordinates corresponds to ISO Standard 4130:1978.

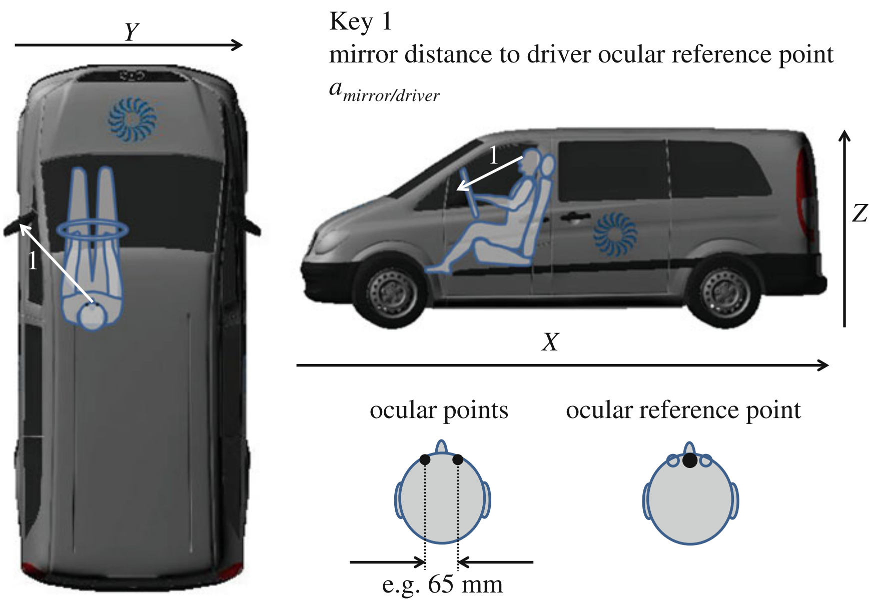


Fig. 20 Coordinate system of the vehicle, distance to the mirror and ORP based on [2]

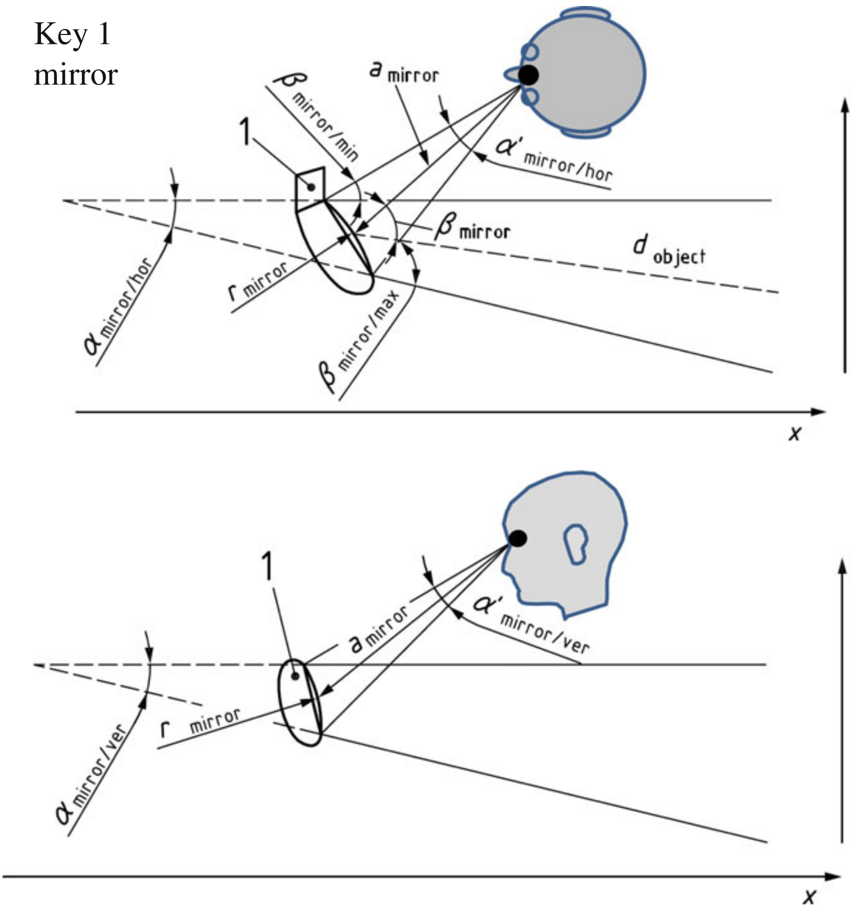


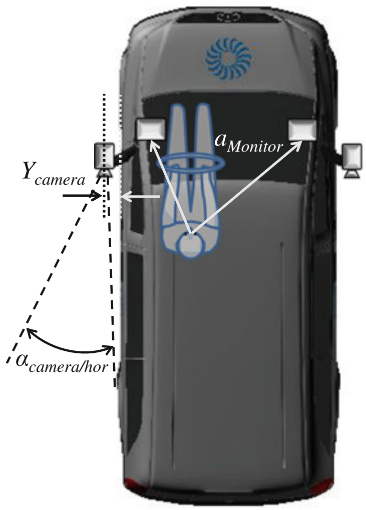
Fig. 21 Mirror horizontal and vertical angular size based on [2]

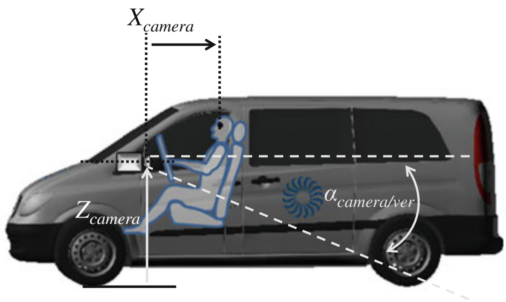
The mirror distance to the driver's ocular reference point is $a_{mirror/driver}$. Figure 21 shows the horizontal and vertical angular size of a mirror. They are denoted $a'_{mirror/hor}$ and $a'_{mirror/ver}$ and are defined in degrees [2]. The mirror viewing angle β_{mirror} is also included in Fig. 21 and can be based on current homologated vehicles (for a Class III example on the driver side) between $\beta_{mirror/driver/min} = 30°$ to $\beta_{mirror/driver/max} = 65°$. The distance between the mirror and a object is denoted d_{object}. The maximum distance between the ocular reference point of the driver to the driver-side mirror can be up to $a_{mirror/driver/max} = 1.2$ m and to the passenger-side mirror up to $a_{mirror/passenger/max} = 1.9$ m in case of Class III [2].

5.2 Parameters for a Class III CMS

In order to present an example with real values for the design parameters, a real vehicle equipped with Class III systems is considered. It is a Mercedes-Benz Vito where CMS cameras are assumed to be mounted at the position of the Class III side-view mirrors and the displays at the A-pillars on the left and right. This is a theoretical example and is used due to the availability of this research vehicle at the Ulm University of Applied Sciences. This research vehicle is used to investigate new vision enhancement technologies such as CMS. Figure 22 shows the camera and display mounting positions on the vehicle, including example values for X, Y and Z positions. The cameras' horizontal and vertical field of view is also shown in Fig. 22. Using the parameter of the camera position, the minimum horizontal field of view of the camera can be calculated with Eqs. 2–4 according to [2]. Equation 2 considers the 1 m/4 m position of the FOV (see Fig. 3) and can be derived as

$$\alpha_{camera/hor\_4\,m} = \left(\operatorname{atan}\left(\frac{1\,m - y_{camera}}{\sqrt{(x_{camera} + 4\,m)^2 + z_{camera}{}^2}} \right) + \operatorname{atan}\left(\frac{y_{camera}}{\sqrt{(x_{camera} + 4\,m)^2 + z_{camera}{}^2}} \right) \right). \tag{2}$$





X_{camera}	0.58 m
Y_{camera}	0.15 m
Z_{camera}	1.41 m
$a_{Monitor}$ (driver side)	0.94 m
$a_{Monitor}$ (passenger side)	1.38 m

Fig. 22 Position of the cameras and the displays

Equation 3 considers the 4 m/20 m position of the same FOV and can be derived as

$$\alpha_{camera/hor\_20\mathrm{m}} = \left(\operatorname{atan}\left(\frac{4\mathrm{m} - y_{camera}}{\sqrt{(x_{camera} + 20\mathrm{m})^2 + z_{camera}{}^2}} \right) + \operatorname{atan}\left(\frac{y_{camera}}{\sqrt{(x_{camera} + 20\mathrm{m})^2 + z_{camera}{}^2}} \right) \right). \tag{3}$$

The minimum cameras horizontal field of view is the maximum value of the two considered positions

$$\alpha_{camera/hor} \geq \max\left[\alpha_{camera/hor\_4\mathrm{m}}, \alpha_{camera/hor\_20\mathrm{m}}\right]. \tag{4}$$

For the given parameters the resulting value is $\alpha_{camera/hor} = 11.85°$.

The minimum cameras' vertical field of view can be derived as

$$\alpha_{camera/ver} = \operatorname{atan}\left(\frac{z_{camera}}{x_{camera} + 4\,\mathrm{m}} \right) \tag{5}$$

and for the parameters used in this example the resulting value is $\alpha_{camera/ver} = 17.11°$. Assuming a camera with a resolution of 1280 × 800 pixels ($N_{camera/hor} \times N_{camera/ver}$ pixels) leads to an aspect ratio of 16:10 (width:height). Considering this aspect ratio, the minimum cameras' horizontal field of view has to be extended to $\alpha_{camera/hor/landscape} = 27.38°$. In the example, a real camera having a horizontal field of view of $\alpha_{camera/hor} = 55°$ is assumed.

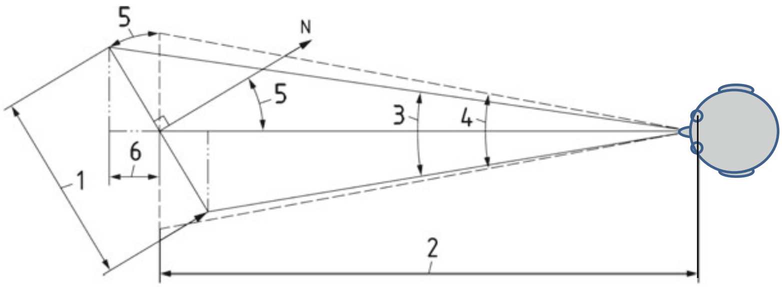
Different parameters of the monitor position need to be considered for the system design. These parameters describe the position and the orientation of the monitor in relation to the driver's eyes. According to ISO 16505:2015, there are different parameters to be considered for the vertical and for the horizontal direction. Figure 23 shows the relevant parameters for the horizontal angular size and Fig. 24 shows the parameters for the vertical situation. The index "*D*" indicates that these values correspond to the monitor design viewing direction, which can be specified by the manufacturer. For the example CMS components presented in this chapter, the optimum design viewing direction would be a perpendicular view.

All lengths, e.g., $a_{monitor/D}$ are defined in m (or mm) and all angles, e.g., $\theta_{monitor/hor/D}$ are defined in degrees.

Based on the value of the monitor horizontal viewing angle $\theta_{monitor/hor/D}$ and the corresponding vertical value $\theta_{monitor/ver/D}$ the monitor viewing inclination angle $\Theta_{monitor}$ can be derived by

$$\Theta_{monitor} = \operatorname{atan}\sqrt{\tan^2\left(\vartheta_{monitor/hor/D}\right) + \tan^2\left(\vartheta_{monitor/ver/D}\right)}. \tag{6}$$

An 8 in. display with dimensions $W_{monitor/hor} = 174$ mm and $H_{monitor/ver} = 104.4$ mm is assumed. The resolution of the display can be assumed to be 800 × 480 pixels ($N_{monitor/hor} \times N_{monitor/ver}$). Table 6 summarizes the CMS parameters for the Class III example.



Key

1 $W_{monitor/hor}$, horizontal size of the monitor, tilted at centre

2 $a_{monitor/D}$

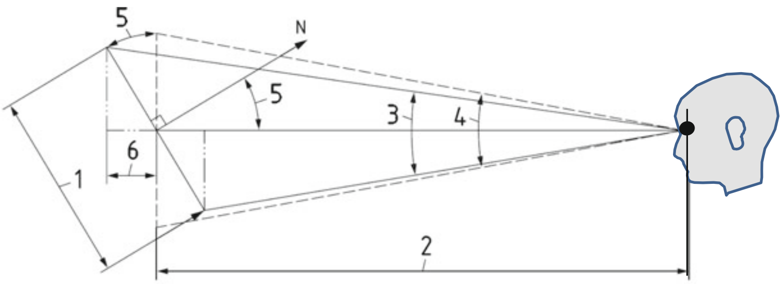
3 $\alpha'_{monitor/hor/D}$

4 $\alpha'_{monitor/hor}$, horizontal angular size of the monitor as seen from $\theta_{monitor/hor} = 0$

5 $\theta_{monitor/hor/D}$

6 $\frac{1}{2} W_{monitor/hor} * \sin\left(\theta_{monitor/hor/D}\right)$

Fig. 23 Monitor design horizontal angular size based on ISO 16505:2015 [2]



Key

1 $H_{monitor/ver}$, vertical size of the monitor, tilted at centre

2 $a_{monitor/D}$

3 $\alpha'_{monitor/ver/D}$

4 $\alpha'_{monitor/ver}$, vertical angular size of the monitor as seen from $\theta_{monitor/ver} = 0$

5 $\theta_{monitor/ver/D}$

6 $\frac{1}{2} H_{monitor/ver} * \sin(\theta_{monitor/ver/D})$

Fig. 24 Monitor design vertical angular size based on ISO 16505:2015 [2]

The system parameters, e.g., the magnification factor show that the system in the example can meet the ISO 16505:2015 requirements. The assumed 8 in. display could be even smaller and still fulfill the requirements. Figure 25 shows the minimum display size for the Class III example on the driver side. According to this diagram, the minimum display size has to be around 6.75 in. in this example. This display represents a larger field of view since the camera has a horizontal field of

Table 6 CMS parameter for a Class III example

	Parameter	Driver side Class III	Passenger side Class III
Camera parameter	Position	Left	Right
	X_{camera}	0.58 m	0.58 m
	Y_{camera}	0.15 m	0.15 m
	Z_{camera}	1.41 m	1.41 m
	$N_{camera/hor}$	1280 pixels	1280 pixels
	$N_{camera/ver}$	800 pixels	800 pixels
	$\alpha_{camera/hor}$	55°	55°
Monitor parameter	Position	Left A-pillar	Right A-pillar
	$a_{Monitor/Driver}$	0.94 m	1.38 m
	$\theta_{monitor/ver/D}$	0°	0°
	$\theta_{monitor/hor/D}$	0°	0°
	$W_{monitor/hor}$	174 mm	174 mm
	$H_{monitor/ver}$	104.4 mm	104.4 mm
	$N_{monitor/hor}$	800 pixels	800 pixels
	$N_{monitor/ver}$	480 pixels	480 pixels
ISO 16505 requirements	$M_{mirror/driver/avg}$	0.31	–
	$M_{mirror/passenger/avg}$	–	0.20
	V_{eye}	0.7	0.7
Calculated values for the camera	$\alpha_{camera/hor/min}$	11.85°	11.85°
	$\alpha_{camera/ver/min}$	17.11°	17.11°
	$\alpha_{camera/hor/landscape}$	27.38°	27.38°
Calculated values for the monitor	$\alpha'_{monitor/hor}$	10.58°	7.21°
	$\alpha'_{monitor/hor/D}$	10.58°	7.21°
	$\alpha'_{monitor/ver}$	6.36°	4.33°
	$\alpha'_{monitor/ver/D}$	6.36°	4.33°
Magnification factor	$M_{system/avg/hor}$	0.37	0.25
	$M_{system/avg/ver}$	0.37	0.25

Fig. 25 Minimum display size for a Class III example at the driver side

Fig. 26 Representation of the Class III FOV at the driver side

view of $\alpha_{camera/hor}$ = 55°. The required magnification factor can also be fulfilled outside the required FOV.

Figure 26 shows the representation of the driver side FOV on the 8 in. display. The required UN R.46 field of view is designated by the pylons in the street. It is obvious that a smaller display could also present the minimum required field of view. As discussed in Sect 1.2, in a conventional mirror the field of view can be expanded through head movement. The example shown in Fig. 26 also covers an expanded FOV and driver has this relatively large field of view in every driving situation without having to move.

5.3 *Architecture and Components of a CMS*

There are different possibilities for designing the architecture of a CMS. The explanations in this subchapter refer to a research prototype CMS that has been designed at Ulm University of Applied Sciences. The CMS research prototype includes two cameras, two displays and an electronic processing unit. The position of the cameras and the monitors along with the relevant parameters are listed in Table 6.

5.3.1 Camera of a CMS

The camera is a high-dynamic-range color CMOS-camera with a resolution of 1280 × 800 pixels. It includes a serializer and transmits the data based on LVDS. It is manufactured by the First Sensor Mobility GmbH in Germany. The camera has an IP68/IP6k9k housing and is designed to be resistant against harsh environmental conditions such as mechanical shocks, vibrations, and hazardous atmospheres [19]. The operating temperature range is relatively high at −40 °C to +105 °C. The optics are watertight and equipped with anti-scratch, anti-reflection, and anti-fog coating. It uses a fixed-focus optics enabling a horizontal field of view (HFoV) 55° ± 2° and can be used for the Class II or III FOV. The same camera is also available with an

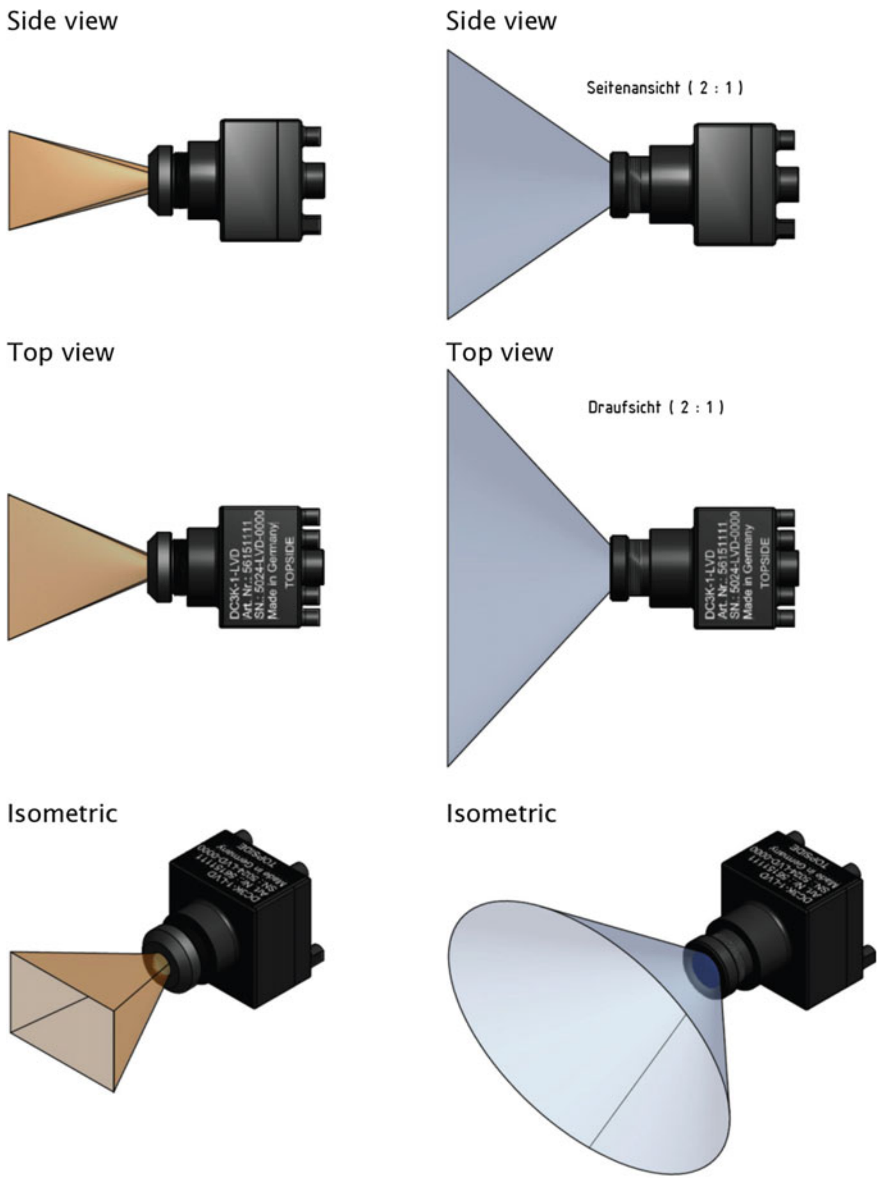


Fig. 27 Camera field of view, left side camera HFoV 55° and right side camera HFoV 100° (picture courtesy of First Sensor Mobility GmbH) [19]

extended HFoV 100° ± 2° and can be used for the Class IV FOV. The field of view of the camera is shown in Fig. 27.

A summary of the technical data of the camera is given in Table 7.

Table 7 Technical data of the First Sensor Mobility GmbH camera [19]

Parameter	Value
Power supply	12 V_{DC}
Current consumption	60 mA @ 12 V_{DC} and 15 fps
	65 mA @ 12 V_{DC} and 20 fps
	80 mA @ 12 V_{DC} and 30 fps
Operating temperature	−40 °C to +105 °C
Sensor type	1/3 in. HDR-CMOS sensor
	Omnivision OV10635
Supported picture size	WXGA (1280 × 800 pixels)
	HD 720p (1280 × 720 pixels)
	WVGA (752 × 480 pixels)
	VGA (640 × 480 pixels)
Sensitivity	Dynamic range: >100 dB (including optics)
	max. S/N ratio: 39 dB
Data interface	10–100 MHz 10/20-bit DC-balanced FPD-link III serializer DS90UB913Q
	10 bit pixel depth up to 100 MHz
	12 bit pixel depth up to 75 MHz
	Bi-directional (I^2C-controller interface at 400 kHz)
Format video stream	Pure transmission of raw signal
General camera parameter	Automated gain control
	Wide dynamic range
	Automated white balance
	Shutter: Electronic rolling shutter
Horizontal field of view	55° ± 2°

The camera has a cubic housing layout with dimensions 26 mm × 26 mm × 28.4 mm. The camera can be connected using a pigtail including an HSD connector. Figure 28 shows the physical dimensions and electrical connection of the camera.

5.3.2 Display of a CMS

The implemented display is an 8 in. (diagonal) color display based on a TFT-LCD (Thin Film Transistor Liquid Crystal Display) unit. It achieves a very high brightness of up to 1200 cd/m^2 and is still discernable in very bright environments. It incorporates a surface treatment with anti-glare and hard-coating. The digital data is transmitted using an LVDS interface. The TFT-LCD module includes the LED driver and backlight components. The dimming function can be controlled using a PWM (pulse-width modulation) signal. A summary of the technical data of the display is given in Table 8.

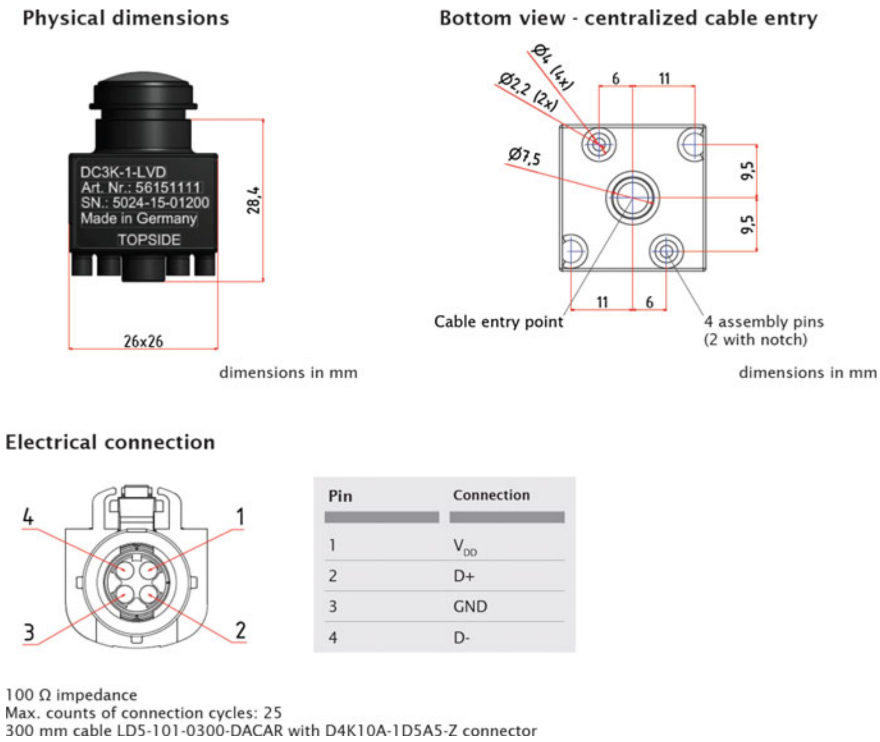


Fig. 28 Camera physical dimensions and electrical connection (picture courtesy of First Sensor Mobility GmbH) [19]

The resolution of the display is lower than that of the camera (see Table 7). It is possible to limit the display to a region of interest of the camera in order to have a pixel by pixel representation. Additional information about the display technology can be found in [20].

5.3.3 Processing Components for a CMS

The CMS architecture of the prototype system at the Ulm University of Applied Sciences is based on using embedded automotive components to fulfill, e.g., the real time requirements and the temperature range. During the CMS research project, different embedded HW and SW variants have been investigated, including FPGA-based designs and System-on-Chip (SoC) designs. The aim was to design a system that is flexible and uses embedded image processing components to support the option of transferring the research results to industrial CMS projects.

Table 8 Technical data of the display

Parameter	Value
Power supply	3.3 V_{DC}
Current consumption	260 mA
Operating temperature	−30 °C to +80 °C
Active screen area (mm)	174.0 (H) × 104.4 (V)
Resolution	800 × 480 pixels
Luminance	1200 cd/m^2
Colors	262 k @ (6 bit/color),
	16.7 M @ (8 bit/color)
Viewing angle	−80° to 80° horizontal
	−80° to 80 vertical
Weight	250 g
Module dimensions	192.0 mm × 122.0 mm × 8.9 mm

One flexible CMS architecture is based on the MB86R12 device, a SoC from Socionext. The MB86R12 includes an ARM® Cortex™ A9 533 MHz processor in combination with an ARM® Neon SIMD engine™ (Single Instruction, Multiple Data engine). In addition it includes a programmable graphic shader GPUs (Graphics Processor Unit) and a chip-internal 2D and 3D image processor. There are 4 independent parallel video ports enabling video stream capture with a resolution of up to 1280 × 720 pixels at 60 fps. Image processing operations such as up-/downscaling or the control of the brightness and contrast can be processed individually for every video channel by using the internal image processor of the chip. A back channel option makes it is possible to control the imager of the camera using I^2C signals. Typical control tasks are the adjustment of the frame rate or the selection of a region of interest. The output of the MB86R12 SoC includes 3 independent display controllers based on Inova Semiconductor's high-speed APIX2® (Automotive/Advanced Pixel Link) bi-directional video interface. The display resolution of every port can be adjusted separately based on a 133 MHz pixel clock and the data rate can be up to 3Gbit/s. Additional information about the video links can be found in chapter "Video Interface Technology".

For a passenger vehicle architecture, this single SoC (MB86R12) can be used to capture the two Class III camera video streams (left and right) and the Class I video stream simultaneously. Independent image processing is possible for every video stream, including different graphical layers for displaying overlay information. The 3 display ports can be used to connect and control the relevant displays for Class I and Class III. Additional combinations of up to 4 cameras are possible. A CAN interface, also included in the SoC, enables the interaction with the vehicle. For commercial vehicle configurations, a Class II and Class IV camera can be used to include different displays for every FOV. Such a configuration is shown in Fig. 29.

A dedicated Graphics Display Controller is required to connect a remote display module to the MB86R12 SoC. Such a device must be able to receive and send data

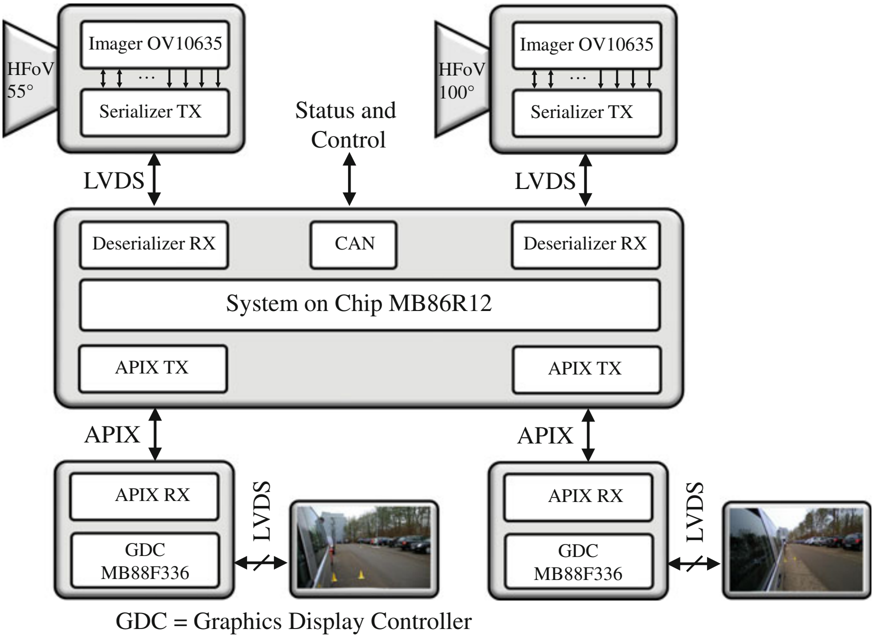


Fig. 29 Architecture of a CMS with two cameras and two displays

based on the APIX2® technology and can be implemented within the display housing. Figure 29 shows the realized architecture, including an embedded MB88F336 Graphics Display Controller from Socionext. This GDC includes an integrated APIX® Phy. as well as additional interfaces such as I^2C and PWM.

The overall system latency of a design such as that shown in Fig. 29 can be <70 ms and fulfills the ISO 16505:2015 requirements in terms of the real time behavior. A CMS can be rated in the ASIL range between QM and ASIL B as explained in chapter "Functional Safety of Camera Monitor Systems". The MB88F336 Graphics Display Controller also includes functions that are required for systems with ASIL B rates. The presented CMS architecture in this subchapter is only one option and serves as an example based on the ISO 16505:2015 requirements. For an even higher quality, it is recommended to use a higher frame rate, e.g., 60 fps and an even higher resolution for the camera and the display. For each project, a dedicated CMS architecture needs to designed and assessed. Additional information on the embedded processing components can be found in this book in chapter "Camera-Monitor-Systems as Solution for the Automotive Market". Additional information on an FPGA-based implementation of the image processing can be found in [21].

6 Conclusion

The new ISO 16505:2015 standard and the latest version of the UN R.46 amendment established an international regulatory framework consisting of a combination of normative and regulatory guidelines. Based on these requirements, it is possible to replace mandatory vehicle mirrors with CMS. This chapter explained the philosophy of the CMS design based on ISO 16505:2015. The standard follows a system-level approach that is independent of different camera, display and processing technologies. The key requirements are the real-time display of a mandatory field of view and the implementation on par with a conventional mirror. The system latency time must therefore be under 200 ms with a video frame rate of at least 30 Hz in daylight conditions. Regarding the CMS resolution, the driver's visual acuity requirements are considered and translated into requirements of observable angular details in the CMS display.

A discussion of the potential benefits showed that a CMS can improve the fuel economy, the aerodynamic, and the aeroacoustics of a vehicle. For a commercial vehicle, a CMS has the potential to improve the fuel consumption by approximately 2 %. One key safety aspect that can be improved with a CMS is vision. It can display an increased field of view and reduce the blind spots in the vehicle environment. Conventional mirrors also distract drivers through glare from the sun or headlights of trailing vehicles. Glare can be significantly reduced with a CMS. Another benefit is the improved vision during night and rain compared to a conventional mirror. Besides the benefits, potential challenges have also been discussed in this chapter, including the integration aspects of the displays within the passenger vehicles and the inability to modify the field of view through head movement.

An example with actual values for a CMS design concluded this chapter. The prototyping CMS includes an HDR camera with a resolution of 1280 × 800 pixels, an 8 in. display, and embedded image processing based on an SoC. CMS technology can form the basis to enable the development of new, advanced driver assistance systems.

References

1. United Nations Regulation No. 46 (2013). E/ECE/324/Rev.1/Add.45/Rev.5, Uniform provisions concerning the approval of devices for indirect vision and of motor vehicles with regard to the installation of these devices, 19 August 2013. http://www.unece.org/fileadmin/DAM/trans/main/wp29/wp29regs/2013/R046r5e.pdf (2013). Accessed 17 Oct 2015
2. International Organization for Standardization, ISO 16505:2015 Road vehicles—ergonomic and performance aspects of camera monitor systems—requirements and test procedures (2015)
3. United Nations Economic and Social Council, ECE/TRANS/WP.29/2015/84, Proposal for Supplement 2 to the 04 series of amendments to Regulation No. 46 (Devices for indirect vision), 27 August 2015. http://www.unece.org/fileadmin/DAM/trans/doc/2015/wp29/ECE-TRANS-WP29-2015-084e.pdf (2015). Accessed 29 Oct 2015

4. Lorenz, L.M.: Entwicklung und Bewertung aufmerksamkeitslenkender Warn—und Informationskonzepte für Fahrerassistenzsysteme. Dissertation. Technischen Universität München (2014)
5. Fischer, G., Much, W., Friederich, M., Schmid, H., Renz, U., Hammerschmied, M.: Innovationen in Sicht. ATZextra **18**(5), 48–55 (2013)
6. Höfer, P., Mößner, A.: Funktion, Sicherheit und Komfort. In: Schütz, T. (ed.) Hucho—Aerodynamik des Automobils, pp. 441–484. Springer, Fachmedien Wiesbaden (2013)
7. European Commission COM (2011) 112 final, A Roadmap for moving to a competitive low carbon economy in 2050 (2011)
8. Kaya, B.: Typgenehmigungsverfahren im Fall neuer Techniken am Beispiel des Volkswagen XL1. http://www.vdtuev.de/sachverstaendigentag/praesentationen/dok_view?oid=515508&vater=515003 (2015) Accessed 18 Oct 2015
9. Schütz, T. (ed.) Hucho—Aerodynamik des Automobils, 6 ed., ATZ/MTZ-Fachbuch Springer Vieweg (2013)
10. Bothe, A.: Analyse dynamischer Sichtsituationen zur ergonomischen Auslegung von Kamera-Monitor-Systemen (KMS) in schweren Nutzfahrzeugen, Dissertation. Technische Universität Darmstadt (2014)
11. Kopp, S., Frank, T.: Nutzfahrzeuge. In: Schütz, T. (ed.) Hucho—Aerodynamik des Automobils, pp. 651–726. Springer Fachmedien Wiesbaden (2013)
12. Devesa A., Indinger, T.: Verbrauchsreduktion an Nutzfahrzeugkombinationen durch aerodynamische Maßnahmen, FAT-Schriftenreihe Nr. 237. Forschungsvereinigung Automobiltechnik e.V., Berlin (2011)
13. Schütz, T., Krüger, L., Lentzen, M.: Luftkräfte und deren Beeinflussung an Personenkraftwagen. In: Schütz, T. (ed.) Hucho—Aerodynamik des Automobils , pp. 177–382. ATZ/MTZ-Fachbuch Springer Vieweg (2013)
14. Neuendorf, R., Zuck, B.: Kühlung und Durchströmung. In: Schütz, T. (ed.) Hucho—Aerodynamik des Automobils, pp. 485–522. Springer Fachmedien Wiesbaden (2013)
15. Spritsparmodelle aus dem Windkanal. http://www.auto-motor-und-sport.de/news/aerodynamik-report-spritsparmodelle-aus-dem-windkanal-4022658.html (2011) Accessed 26 Oct 2015
16. We fight for every third decimal place, Daimler AG. https://www.daimler.com/dccom/0-5-1629048-1-1629834-1-0-0-0-0-0-0-7165-0-0-0-0-0-0-0.html (2015) Accessed 26 Oct 2015
17. Farmer, C.M.: Crash Avoidance Potential of Five Vehicle Technologies, Insurance Institute for Highway Safety (IIHS) (2008)
18. European Commission Directorate-General For Energy And Transport, Fitting Blind-Spot Mirrors On Existing TruckS, 12 April 2006. http://ec.europa.eu/transport/road_safety/consultations/doc/2006_05_19_safer_trucks_consultation_paper.pdf (2006) Accessed 30 Oct 2015
19. First Sensor Mobility GmbH, HDR-CMOS megapixel imager sensor device with LVDS interface. http://www.first-sensor.com/cms/upload/datasheets/DS_Standard-Blue_eagle_DC3K-1-LVD_E_9999-09.pdf (2015) Accessed 13 Nov 2015
20. Chen, J., Cranton, W., Fihn, M.: Handbook of Visual Display Technology, Springer (2012)
21. Bailey, D.G.: Design for Embedded Image Processing on FPGAs, Wiley-IEEE Press (2011)

Standardization and Vehicle Regulation Aspects of Camera Monitor Systems

Matthias Esser

Abstract Inside and outside rearview mirrors provide drivers of motor vehicles a field of vision that is relevant for safe driving behavior in several driving situations. This is one of the reasons why mirrors are a legal requirement almost everywhere in the world. UN Regulation No. 46 is the first regulation that permits all mandatory mirrors for passenger cars, commercial vehicles and buses to be replaced by camera monitor systems (CMS). The amended regulation is based on the new international standard ISO 16505 "Ergonomic and performance aspects of Camera Monitor Systems—Requirements and test procedures". This chapter describes the relevance and impact of vehicle regulations and standards on vehicle development. On this basis, the development process of standards and regulations is examined using the examples of ISO 16505 and UN Regulation No. 46. It then proceeds to analyze today's regulatory constraints with regard to CMS for passenger cars and highlights the basic boundary conditions which have to be considered within the development process to ensure product compliance with the relevant regulations. The chapter concludes with an overview of how future activities could contribute to regulation and standardization of CMS.

Keywords Camera monitor system · Replacement of rearview mirrors · Standardization · Developing standards ISO 16505 · Vehicle regulations · Developing vehicle regulations · Vehicle certification · UN Regulation No. 46 · Requirements for camera monitor systems

List of Acronyms

AA	Arbeitsausschuss (Technical Committee)
AC	Administrative Committee
ACEA	European Automobile Manufacturers' Association
AIS	Automotive Industry Standard
AK	Arbeitskreis (Working Group)

M. Esser (✉)
Mercedes-Benz Cars Development, Certification and Regulatory Affairs Active Safety, Daimler AG, Sindelfingen, Germany
e-mail: matthias.esser@daimler.com

A. Terzis (ed.), *Handbook of Camera Monitor Systems*,
Augmented Vision and Reality 5, DOI 10.1007/978-3-319-29611-1_2

ANSI	American National Standards Institute
BASt	Bundesanstalt für Straßenwesen (Federal Highway Research Institute of the Republic of Germany)
CD	Committee Draft
CEN	European Committee for Standardization
CENELEC	European Committee for Electrotechnical Standardization
CFR	Code of Federal Regulations
CLEPA	European Association of Automotive Suppliers
CMS	Camera Monitor System
CMVR	Central Motor Vehicle Rules
CMVSS	Canadian Motor Vehicle Safety Standard
CS	ISO Central Secretariat
DIN	Deutsches Institut für Normung (German Institute for Standardization)
DIS	Draft International Standard
DKE	Deutsche Kommission Elektrotechnik Elektronik Informationstechnik in DIN und VDE (German Commission for Electrical, Electronic and Information Technologies of DIN and VDE)
EC	European Community
ECE	Economic Commission for Europe
EMC	Electromagnetic Compatibility
EN	European Standard
ETSI	European Telecommunications Standards Institute
EU	European Union
FDIS	Final Draft International Standard
FMVSS	Federal Motor Vehicle Safety Standard
FTA	Free Trade Agreement
GB	National Standard
GEB	Group of Experts on Noise
GEE	Group of Experts on Lighting and Light-Signaling
GEPE	Group of Experts on Pollution and Energy
GERF	Group of Experts on Brakes and Running Gear
GESG	Group of Experts on General Safety
GESP	Group of Experts on Passive Safety
GR	Groupe des Rapporteurs (Group of Experts)
GRB	Working Party on Vehicle Noise
GRE	Working Party on Lighting and Light-Signaling
GRPE	Working Party on Pollution and Energy
GRPS	Working Party on Passive Safety
GRRF	Working Party on Brakes and Running Gear
GRSG	Working Party on General Safety Provisions
HMI	Human Machine Interface
IEC	International Electrotechnical Commission

IG	Informal Group
IS	International Standard
ISO	International Organization for Standardization
ITC	Inland Transport Committee
ITU	International Telecommunication Union
JAMA	Japan Automobile Manufacturers' Association
JAPIA	Japan Auto Parts Industries Association
KBA	Kraftfahrtbundesamt (German Federal Motor Transport Authority)
KMVSS	Korean Motor Vehicle Safety Standards
MEMA	Motor & Equipment Manufacturers Association in the USA
MLIT	Ministry of Land, Infrastructure and Transport
MLTM	Ministry of Land, Transport and Maritime Affairs
MOLIT	Ministry of Land, Infrastructure and Transport
MTF	Modulation Transfer Function
NA	Normenausschuss Automobiltechnik (Automotive Standards Committee)
NCAP	New Car Assessment Program
NGO	Non-Governmental Organization
NHTSA	National Highway Traffic Safety Administration
NP	New Work Item Proposal
OICA	Organisation Internationale des Constructeurs d'Automobiles (International Organization of Motor Vehicle Manufacturers)
PRC	People's Republic of China
PWI	Preliminary Work Item
R.E.3.	Consolidated Resolution on the Construction of Vehicles
SAE	Society of Automotive Engineers
SC	Subcommittee
TC	Technical Committee
TCMV	Technical Committee Motor Vehicles
TS	Technical Specification
UN	United Nations
UNECA	United Nations Economic Commission for Africa
UNECE	United Nations Economic Commission for Europe
UNECLAC	United Nations Economic Commission for Latin America and the Caribbean
UNECOSOC	United Nations Economic and Social Council
UNESCAP	United Nations Economic Commission for Asia and the Pacific
UNESCWA	United Nations Economic Commission for Western Asia
UNO	United Nations Organization
UN-R	UN Regulation (annexed to the 1958 Agreement)
USA	United States of America

VDA	Verband der Automobilindustrie (German Automobile Industry Association)
VDA NA	VDA Normenausschuss Automobiltechnik (VDA Automotive Standards Committee)
VDE	Verband der Elektrotechnik und Elektronik und Informationstechnik (Association for Electrical, Electronic and Information Technologies)
WG	Working Group
WP.29	UNECE World Forum for Harmonization of Vehicle Regulation (formerly: UNECE Working Party 29)

1 Standards and Regulations

1.1 Requirements for Vehicles

This section introduces requirements necessary for incorporation into the early phases of the development of a vehicle (see Fig. 1).

A vehicle has to comply with vehicle regulations. This is verified by the manufacturer within the vehicle certification process. There are several types of different requirements the manufacturer needs to address in appropriate form. These include

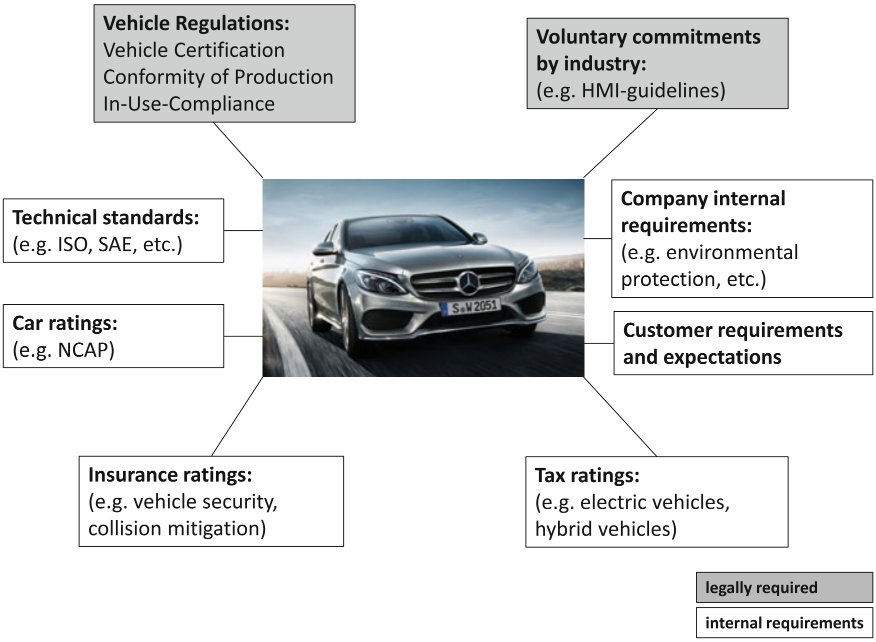


Fig. 1 Requirements for vehicles (© Daimler AG)

technical requirements such as standards, defining the state-of-the-art development. Other examples are customer and competition requirements, including tax ratings that allow tax savings for customers, if the vehicle fulfills certain criteria. The same applies to insurance ratings which have an impact on the insurance ranking of the vehicle (insurance premium rate). The following sections focus on vehicle regulations and standards, as both of these have a significant impact on camera monitor systems.

1.2 Standards

To understand the importance of ISO 16505, this section provides an overview of what a technical standard is, including its relevance for and impact on vehicle development.

"A standard is a document that provides requirements, specifications, guidelines or characteristics that can be used consistently to ensure that materials, products, processes and services are fit for their purpose" (source [5]).

Standards are established under defined rules (see e.g. all parts of DIN 820 series), either in-house (e.g. Mercedes-Benz Standards), within associations (e.g. VDA, SAE), on the national level (e.g. DIN, ANSI), the European level (EN) or the international level (e.g. ISO, IEC) (see Fig. 2).

Revision of Standards

Standards document the state of the art and are revised on a regular basis (the exact timing depends on the rules of the individual organization). The content of ISO standards is reviewed at least every 5 years to reflect the state-of-the-art

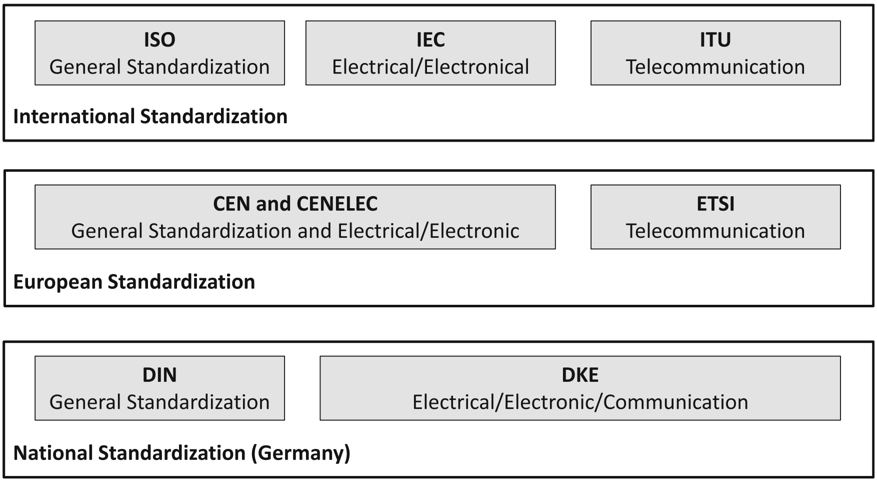


Fig. 2 Some standardization bodies in which VDA NA is involved, *source* [40]

development. The technical experts of the responsible working group evaluate whether a particular standard needs a revision or if it perhaps even needs to be withdrawn (see [41]).

Legal Relevance

The application of standards is not directly legally binding. However, they are used in court (e.g. in product liability cases) to determine current technological state of the art. The current level of scientific and technical knowledge has to be considered for product development. If the product deviates from applicable standards, the manufacturer should prepare a justification (engineering judgment) of why it is still in line with current state-of-the-art technology. Especially for new technologies and innovations, even drafts of standards may be relevant during the product development.

Standards may become part of the legal requirements. As soon as a standard is referred to in a regulation, compliance is mandatory. For legal compliance, the exact standard version mentioned in the regulation has to be used. Therefore, the date of issue mentioned in the regulation has to be thoroughly checked.

With the publication of the new international standard ISO 16505 "Ergonomic and performance aspects of Camera Monitor Systems—Requirements and test procedures" in April 2015 (see [9]), a state-of-the-art replacement of mirrors by CMS was defined for the first time ever.

1.3 Developing ISO Standards

This section explores how ISO standards are developed. Such background information is important to understand the development process of ISO 16505, which is outlined subsequently in Sect. 1.4.

ISO standards are developed by technical expert working groups (WGs), which are organized into Technical Committees (TCs) and Sub Committees (SCs) (see [7], p. 15). There are more than 250 TCs in total and each of them is responsible for specific subjects (see [6]).

The technical experts are sent by ISO's member countries and are e.g. representatives of the industry, NGOs, governments, Technical Services and other stakeholders (see [6]). The ISO members are countries (individuals or companies cannot become ISO members) and each country can only be represented by a single institution (which is typically the national standard body, e.g. DIN for Germany). Table 1 shows the different types of ISO full membership.

Timeline for ISO Standards

ISO provides three different development tracks to develop an ISO standard (see Table 2). Figure 3 shows the process for creating an ISO standard. The important milestones on the roadmap towards an International Standard (IS) are: Proposal new

Table 1 ISO full membership, see [7], pp. 9 and 19

	Full member (member bodies) (can chose whether to participate in any TC as a P- or O- member)	
	P-member	O-member
Participation	Actively participating and influencing standardization work in TCs	Observe TC's work
Comments	Are expected to submit comments	Can submit comments
Votes	Are expected to cast ballots	Can cast ballots
Can adopt ISO standards nationally	Yes	
Can sell ISO standards nationally	Yes	

Table 2 Options for development of an ISO standard, see [7], p. 29

Option	Development track	Time to DIS (months)	Time to publication (months)
1	Accelerated	12	24
2	Standard	24	36
3	Extended	36	48

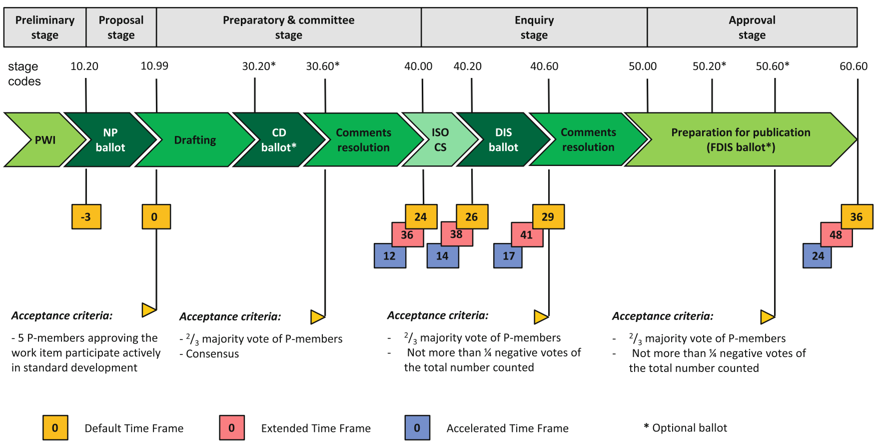


Fig. 3 Developing ISO standards, *source* [40], p. 15

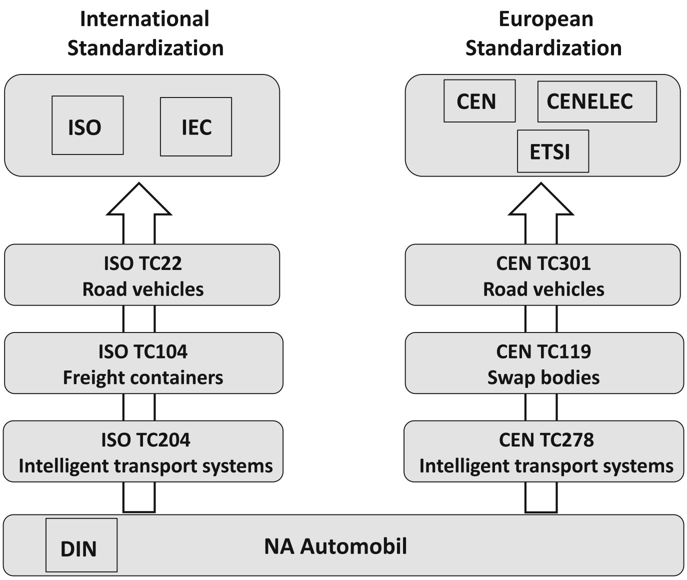


Fig. 4 NA's activities in automotive standardization, *source* [40], p. 5

project (NP), Committee Draft (CD), Draft International Standard (DIS) and Final Draft International Standard (FDIS). Each milestone has an international stage code (see [8]).

Representation of National Interests Exemplified by Germany

The members of an ISO Working Group usually prepare national comments and positions on an ISO standards project within national mirror working groups. These groups belong to the national standards organization (e.g. DIN for Germany).

In Germany, the VDA "Normenausschuss Automobiltechnik" (NA) acts in place of DIN within TCs of ISO and CEN dealing with motor vehicles (see Fig. 4). The NA is organized in several national subcommittees similar to the organization of the TC for road vehicles in ISO and keeps the members in Germany informed about current standardization activities.

1.4 Developing ISO Standard 16505

Figure 5 shows the development timeline of ISO 16505. After the New Work Item Proposal was submitted, it took five years before the International Standard ISO 16505 was published. After registration of the Final Draft International (FDIS), the project faced a delay of approximately 6 months before the FDIS ballot was started because several formal and editorial questions had to be clarified.

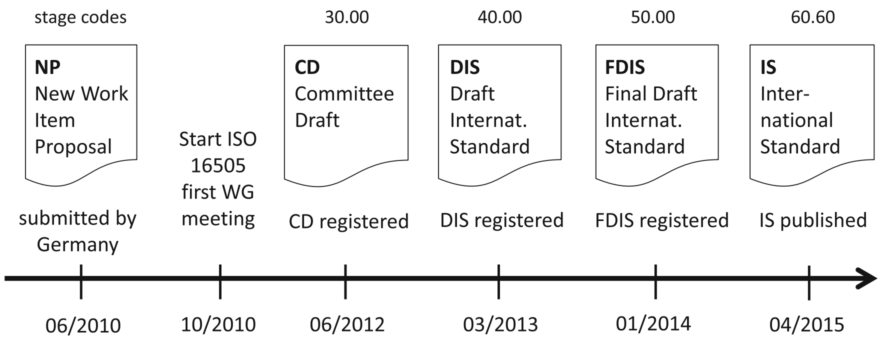


Fig. 5 Timeline for development of ISO 16505

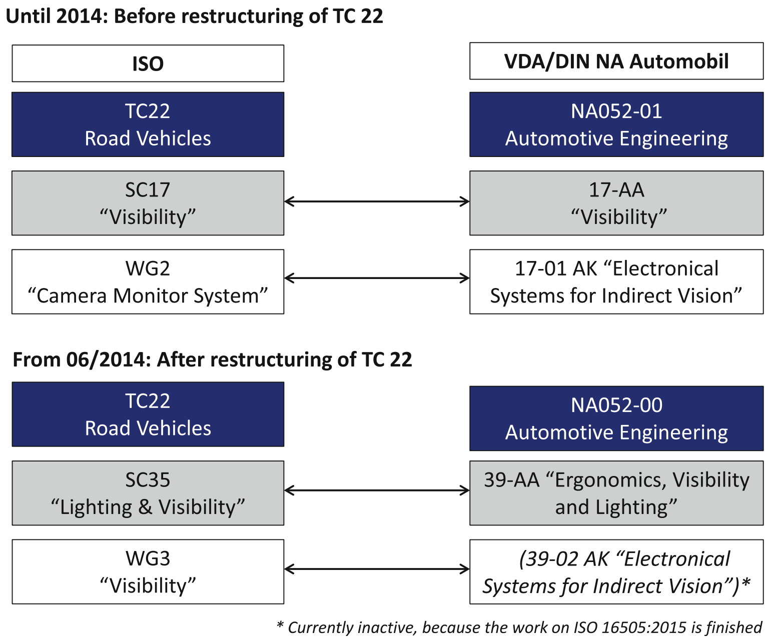


Fig. 6 International and national committee structure for developing ISO 16505

Figure 6 shows the international committee structure in ISO for developing ISO 16505. The expert group was TC22/SC17/WG2 "Camera Monitor Systems". After WG2 had finished its work on ISO 16505 in 2014, TC22 was restructured. As a consequence, the subject ISO 16505 was attributed to TC22/SC35/WG3 "Visibility" within this new structure.

Using the example of Germany, Fig. 6 shows how the international standardization work of ISO is mirrored by national working groups. As explained in

Sect. 1.3, Germany's interests are represented within ISO by the VDA NA Automobil. For camera monitor systems, the working group "Electronic Systems for Indirect Vision" is in charge of ISO 16505.

1.5 Vehicle Regulations and Certification

To understand the importance of UN Regulation No. 46, this section gives an overview of what a vehicle regulation is, including its relevance and impact for the vehicle development.

A technical standard like ISO is not sufficient for CMS to become officially approved for the use on public roads. Therefore, a legal framework is needed which defines the minimum requirements when rearview mirrors are replaced by CMS.

Legal Relevance

Within the framework of vehicle certification, a vehicle manufacturer demonstrates that a vehicle or a system installed in the vehicle complies with the requirements of the relevant regulation. The vehicle certification is different from a certification of a company and its processes, such as certification according to ISO-9000ff or TS 16949, because vehicle certification is mandatory by law whereas certification according to ISO-9000ff is typically a mutually agreed part of an intercompany relationship (Quality Management). Table 3 gives an overview of comparison between standards and vehicle regulations.

There are several hundred vehicle regulations worldwide, prescribing minimum requirements with regard to different functions or systems of a vehicle. Vehicle regulations basically address environmental topics (e.g. exhaust emissions, fuel consumption, noise, etc.), passive safety topics (e.g. restraint systems, different crash configurations, pedestrian protection, etc.) and active safety topics (e.g. braking and steering systems, lighting, rearview mirrors, etc.).

These requirements are not necessarily harmonized internationally. In addition, the certification procedures for the different markets can be different, thus making it necessary to conduct independent certification activities for individual markets.

Table 3 Comparison between standards and regulations

	Standards	Vehicle safety regulations
Purpose	Define the state of the art "to ensure that materials, products, processes and services are fit for their purpose" (source [5])	Define minimum requirements to ensure road traffic safety
Content	Requirements, recommendations, design/product specifications, test procedures	Requirements and test procedures, administrative provisions
Issuer	Individual standards organization	Authorities
Legal relevance	No (Yes, in case a standard is referred to in a regulation)	Yes

Revision of Vehicle Regulations
The creation of new regulations and the further development of existing regulations can have several triggers and motivations. These include enhancement of safety, reduction of environmental pollution, alignment of the regulations with the latest technologies, and harmonization of the regulations, including consideration of regional differences. For CMS, an update of UN Regulation No. 46 was necessary to align with the new technology and to permit it as a compliance option.

1.6 Developing UN Regulations

This section explores how UN Regulations are developed. Such background information is important to understand the development process of UN Regulation No. 46 towards CMS, which is outlined subsequently in Sect. 1.7.

UN Regulations (formerly ECE-Regulations) are significantly important in an international context. They are created by working groups of the so-called United Nations Economic Commission for Europe (UNECE). As shown by Fig. 7, the UNECE is a subcommittee of the United Nations Organization (see [24,39]).

The legal framework for the adoption and enforcement of UN Regulations is known as the 1958 Agreement and was established in 1958. With its latest revision from 1995 (see [37]), non-UNECE countries are also allowed to participate in the vehicle regulations work. As of September 2015, the 1958 Agreement had 52 signatories, including the European Union and its member states, as well as Australia, Japan, Russia, Turkey and many others. A complete list of the current signatories is shown in Annex A.

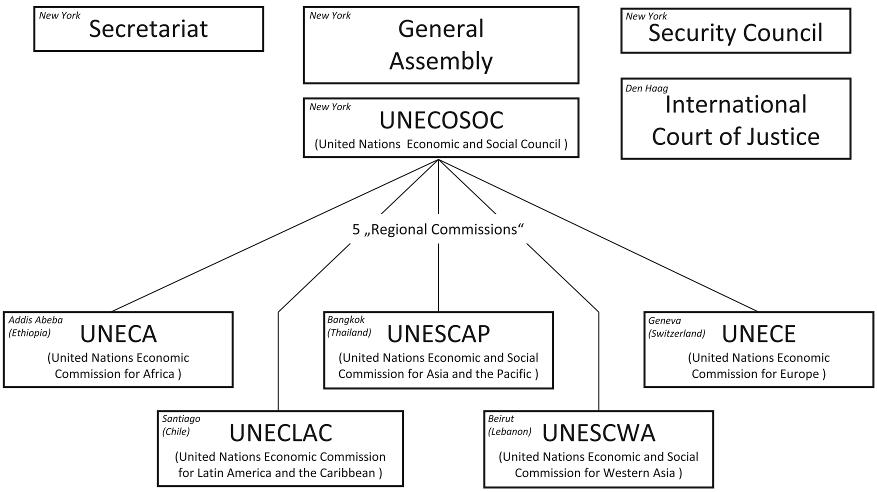


Fig. 7 Integration of UNECE in UNO

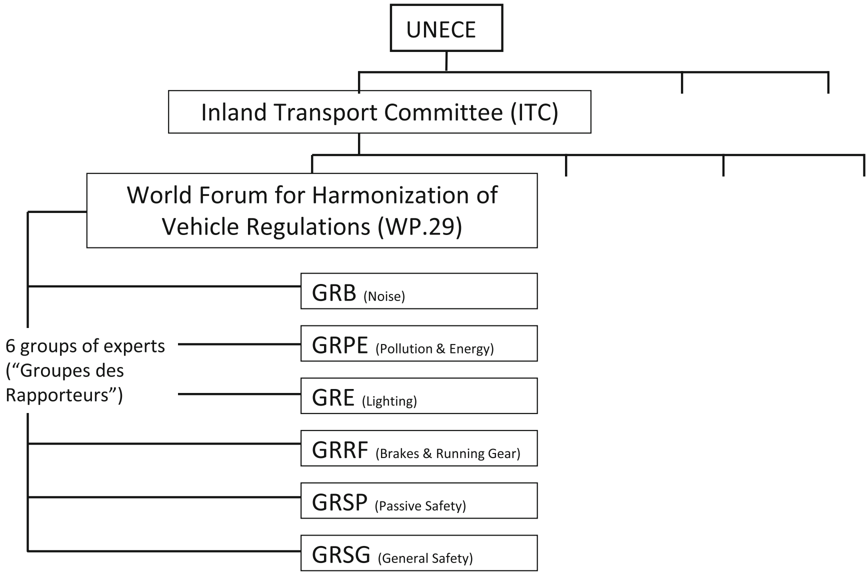


Fig. 8 Organization of WP.29, *source* Based on [42], p. 10

Within UNECE, Working Party 29 (WP. 29) is called the World Forum for Harmonization of Vehicle Regulations. The functional work is organized by six groups of experts (Groupe des Rapporteurs, GR-Groups) (see Fig. 8). The results of those six groups are presented at WP.29 sessions, which are held three times per year. WP.29 reviews and discusses the proposals that the expert groups submitted.

It is up to each Contracting Party to decide whether or not to adopt a UN Regulation and apply it on a national level. However, once a certain UN Regulation is adopted, approvals for the corresponding regulation granted by other Contracting Parties must automatically be accepted ("mutual recognition"). A reason why Contracting Parties may decide not to adopt a UN Regulation is usually that they have established their own national requirements that are not entirely covered by the respective UN Regulation. If they do not adopt it, mutual recognition is not necessary.

Participation in WP.29

Any member country of the United Nations can participate as a full member or in a consultative capacity in WP.29 and become a Contracting Party to the 1958 Agreements (see [42], p. 3).

NGOs can participate only in a consultative capacity (see [42], p. 5). Vehicle manufacturers are represented by OICA (Organisation Internationale des Constructeurs d'Automobile) at WP.29 and its GR-Groups. Vehicle component suppliers are represented by CLEPA (Comité de liaison européen des fabricants d'équipements et de pièces automobiles), MEMA (Motor & Equipment Manufacturers Association in the USA) and JAPIA (Japan Auto Parts Industries Association).

OICA

OICA is the only vehicle manufacturers' association which is accredited by the United Nations. It is not an association of individual vehicle manufacturers. Rather, its members consist of automotive associations of single countries such as Alliance (USA), JAMA (Japan) and VDA (Germany). In 2015, there were 38 members (see [19]).

The task of the OICA Technical Committee consists of coordinating the technical activities of its member associations and in particular, of serving as the voice of the associations in WP.29 as well as its GR groups. The TC is supported by its 6 expert groups (Groupe des Experts, GE-Groups, see Fig. 9), which are mirrored to the 6 working groups (Groupe des Rapporteurs, GR-Groups) of WP.29 (see [20]).

Timeline for UN Regulations

Unlike under the ISO standards, there is no defined timeframe for developing UN Regulations under the 1958 Agreement. Basically, UN Regulations (new regulations and amendments of existing regulations) are developed by the responsible GR working group. The time the GR groups require to prepare a first working draft depends on the complexity of the topic and the challenge to align potentially different interests/opinions within the group. Figure 10 shows an exemplary timeline for the development of UN Regulations. For new and complex topics, GR groups normally install the so-called Informal Groups. Participants in such Informal Groups are typically representatives of Contracting Parties, technical experts of NGOs, like OICA and CLEPA and as well of Technical Services. The GR groups

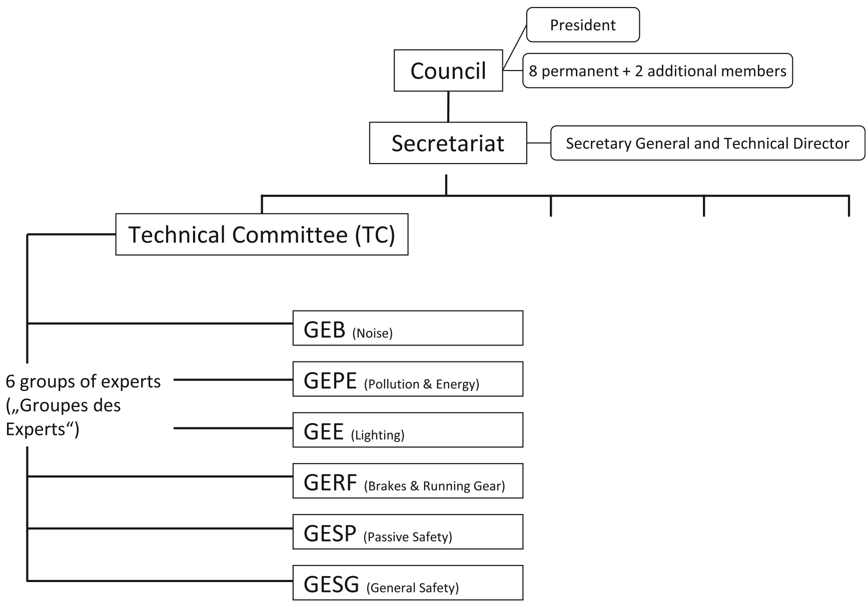


Fig. 9 Organization of OICA's operation and activities with focus on WP.29

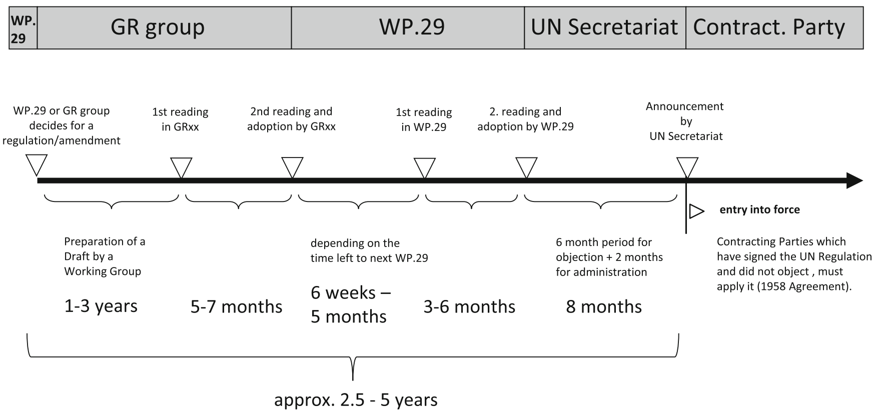


Fig. 10 Exemplary timeline for UN Regulations

present their working results to WP.29. After adoption, WP.29 forwards the regulation to the UN Secretariat in New York.

The Secretary-General of the UN will officially publish the adopted regulation following a 6 months period during which the Contracting Parties have the right to object (see [37], Article 7). If no objection is raised, the regulation is accepted. After the 6 months objection period, it takes about another 2 months for the UN Secretariat to formalize the document, before the Secretary-General of the UN officially announces the regulation and it thereby enters into force.

1.7 *Developing UN Regulation No. 46 Towards CMS*

Before the amendment of UN Regulation No. 46 (see [32]), passenger cars had to be equipped with mirrors, in fact with one inside mirror (Class I) and two outside mirrors (Class III); see Fig. 11. The replacement of those mandatory mirrors by CMS was not permitted.

The mandatory mirrors for commercial vehicles were also not permitted to be replaced by CMS, except for the close-proximity mirror (Class V) and the front mirror (Class VI); see Fig. 11. Since 2005, it has been permitted to replace Class V and VI mirrors by CMS which are only used for low speed applications such as maneuvering. Due to the size and installation position of Class V and VI mirrors, no significant fuel savings could be expected by their replacement. For that reason, Class V and VI CMS have only had a limited market penetration (mainly Class VI front camera applications). Figure 12 shows an example for a Class VI CMS.

The development of UN Regulation No. 46 towards including camera monitor systems for devices of Class I, II, III and IV started in 2009, when WP.29 approved the establishment of GRSG Informal Group Camera Monitor Systems (IGCMS) (see Fig. 13).

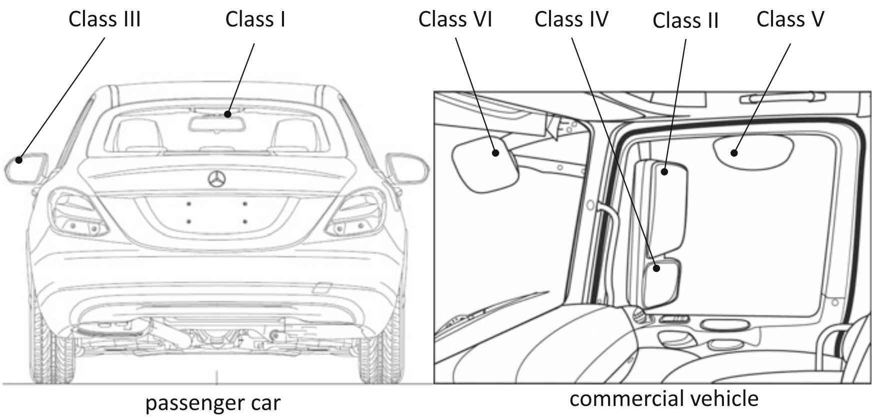


Fig. 11 Mirror classes in UN Regulation No. 46 (© Daimler AG)

Fig. 12 Class VI Camera Monitor System (© MEKRA Lang GmbH & Co. KG)

In light of the discussions it became obvious that the development of technical requirements for such a new technology was complicated, and would therefore need further international expertise. Progress was difficult due to the fact that no regulation or standard yet existed which would serve as a basis to develop technical requirements and test procedures. Due to the completely different use cases (especially vehicle speed), the requirements for Class V and VI mirrors were not suitable and sufficient for the other mirror classes. In 2010, the IGCMS ended up presenting a first draft for amendment of Regulation No. 46 which would permit all mirrors to be replaced by a camera monitor system (see [31]). However, regarding the technical requirements for camera monitor systems of the classes other than V and VI, this first draft referred to a new ISO standard (which had not yet been developed).

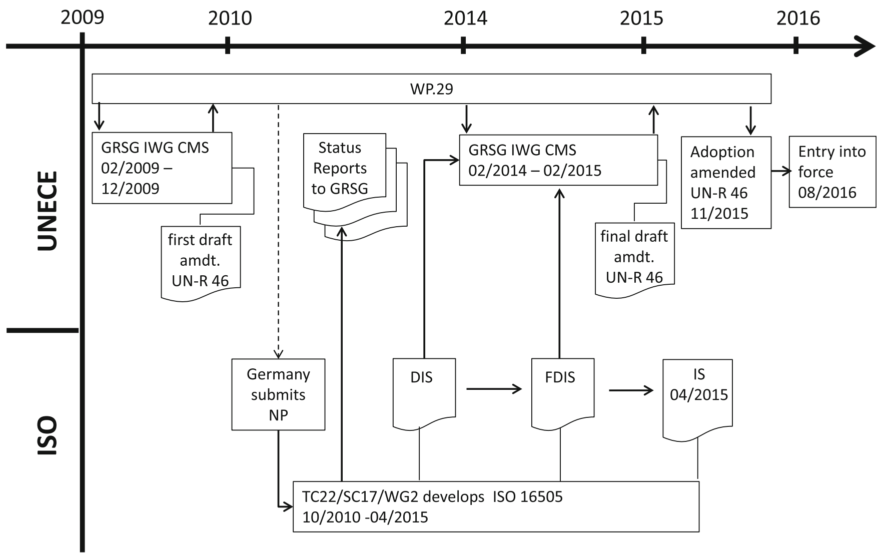


Fig. 13 Process for amendment of UN Regulation No. 46

Based on this outcome, Germany triggered a New Work Item Proposal for ISO in June 2010 (see Figs. 5 and 13). It was approved, and TC22/SC17/WG2 held its initial meeting in November 2010. In the following working period, the chair of the ISO working group regularly reported the current status of the ISO activities at GRSG sessions. During the preparation of FDIS, WP.29 approved the reestablishment of the GRSG Informal Group Camera Monitor Systems (IGCMS). The IGCMS restarted its work by reviewing the technical requirements of the DIS. Meanwhile the IGCMS was kept informed on the outcome of the ongoing ISO work on the FDIS, which ran in parallel during the period from 2014 to 2015. In the further course of discussions, it became obvious that a simple reference to the new ISO standard would not be sufficient. While some provisions were regarded as inappropriate for type-approval purposes (e.g. how to deal with recommendations which may lead to different interpretations), others were determined to be inacceptable, or rather insufficient (e.g. point light sources). The first draft amendment of the IGCMS from 2009 therefore needed further modifications. The group managed to provide a final draft proposal to amend UN Regulation No. 46 to GRSG, which was adopted in May 2015 (see [33]). This proposal was then adopted by WP.29 in November 2015 (see [34]). The amended regulation will enter into force around August 2016.

UN Regulation No. 46 is not only accepted by the European member states. As of September 2015, there are 42 UNECE Contracting Parties which have adopted this regulation. Annex B shows the Contracting Parties that signed UN Regulation No. 46 and where CMS can be installed in the future.

The details of the amended UN Regulation No. 46 are explained in Sect. 3.

2 Regulatory Situation for CMS in Regulations Other Than UN-R 46

2.1 General Aspects

Many of today's vehicles are already equipped with several camera systems (see Fig. 14). Rearview cameras, which are installed at the rear part of the vehicle, support the driver during parking and maneuvering scenarios. 360°-visibility systems are composed of several single cameras installed around the vehicle and whose single images are fused to provide the driver with an enhanced field of vision during low speed applications such as parking and maneuvering. Mono- and stereo-cameras, installed in the upper section behind the windshield, are used to detect traffic signs, lane markings and objects to assist the driver in exercising the longitudinal and lateral control of the vehicle. Furthermore, night view systems with infrared cameras have been on the market for several years. The installation of such driver assistance systems is not required by technical vehicle regulations and the camera systems which form a part of it have not been covered by the scope of regulations so far (except in the USA where the installation of rearview cameras will become mandatory for passenger cars and light trucks as of 1 May 2016, see Sect. 2.5).

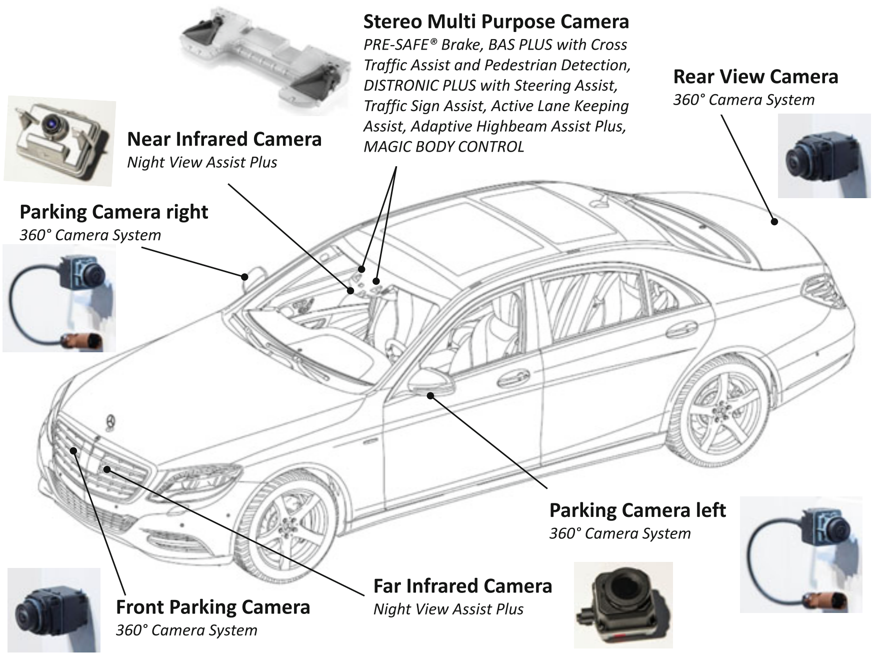


Fig. 14 Cameras in the Mercedes-Benz S-class (V222) (© Daimler AG)

In contrast, camera monitor systems covered by ISO 16505 aim to replace the inside and outside mounted mirrors, which are legally required.

The preceding Sect. 1.7 focused specifically on the regulatory situation with respect to the amended UN Regulation No. 46. The following will analyze to what extent a replacement of mandatory mirrors would be in line with the current vehicle regulations in major markets that have not adopted UN Regulation No. 46.

2.2 Camera Monitor Systems in China

Regulation for indirect vision: GB 15084-2013 (devices for indirect vision) (see [21]).

The content of the regulation is based on the requirements and test methods as defined in UN Regulation No. 46. For that reason, mirrors are also mandatory in China and their replacement by CMS (except for Class V and VI devices) is not yet addressed.

Regulatory Activities to Enable CMS

In the past, GB 15084 was updated within a certain period in order to incorporate the latest amendments of UN Regulation No. 46. The last amendment is from 2013. A further update in the near future does not appear to be a priority for the authority. Therefore, an update and upgrade of GB 15084 to allow camera monitor systems is not to be expected in the near future.

2.3 Camera Monitor Systems in Japan

Regulation for rearview mirrors: Safety Regulation Art. 44 (see [12]).

Mirrors must be present. Their replacement by CMS is not yet addressed.

Since 2005, Japan's vehicle requirements include a unique field of vision (see [11]): A cylinder (height: 1 m; diameter: 0.3 m) must be visible to the driver either by direct or indirect vision when positioned adjacent to the immediate front of the vehicle and to the passenger side. Vehicles with high front fenders have to be equipped with an additional fender mirror (see Fig. 15) in order to fulfill this requirement, because the cylinder would not be visible at some positions neither by direct nor by indirect vision through the outside mirror on the passenger side. The regulation permits the use of camera monitor systems to detect the cylinder. Such an additional fender mirror may therefore be replaced by a CMS.

Regulatory Activities to Enable CMS

Japan as a Contracting Party of the 1958 Agreement intends to adopt UN Regulation No. 46 in the near future. On that basis, the approval of CMS in place of rearview mirrors would become possible in Japan.

Fig. 15 Fender mirror of Mercedes-Benz G-Class (G463), right hand drive for Japan (© Daimler AG)

2.4 Camera Monitor Systems in South Korea

Regulation for rearview mirrors: KVMSS Art. 50 (see [14]).

Rearview mirrors must be present. Their replacement by CMS is not yet addressed.

Since July 2011, Korea and the EU have entered into a free trade agreement (FTA) which also includes motor vehicles and parts. As a proof of compliance with Korean obligations under the FTA, Korea accepts some regulation approvals according to UN Regulations (and vice versa). Within the last three years, the EU and Korea have increasingly harmonized regulations to UN Regulations. In case of Article 50 (driver's visibility through rearview mirrors), the FTA prescribes that UN Regulation No. 46 is accepted as equivalent to the vehicle certification. It would therefore be permissible for vehicles which are exclusively imported from the EU to Korea to be equipped with CMS instead of mirrors as long as the CMS has an approval according to UN Regulation No. 46.

Regulatory Activities to Enable CMS

As of the completion of this chapter, Korean regulatory activities had yet to permit the replacement of rearview mirrors by CMS.

2.5 Camera Monitor Systems in the USA and in Canada

Regulation for rearview mirrors: FMVSS 111 (see [16]), CMVSS 111.

Rearview mirrors must be present. Their replacement by CMS is not yet addressed.

For passenger cars and light trucks (up to 4536 kg) the installation of rearview video systems (camera and monitor) will become mandatory in order to survey an area of 3 m width and 6 m length behind the vehicle (see [17]). By introducing this regulation, the National Highway Traffic Safety Administration (NHTSA) aims to address crashes with pedestrians in the area immediately behind the vehicle while backing up. The introduction period lasts from 1 May 2016 until 1 May 2018 (100 % compliance).

Regulatory Activities to Enable CMS

In March 2014, the Alliance of Automobile Manufacturers and Tesla Motors, Inc. petitioned the agency (NHTSA) to update and upgrade FMVSS 111 to permit camera-based systems as a compliance option (see [1]). As the agency is expecting safety benefits in order to permit CMS, additional safety benefit studies (see e.g. [3]) have been submitted by the Alliance. As of the completion of this chapter, the dialogue between the automotive industry and NHTSA is still ongoing. As this topic is not included in NHTSA's rulemaking plan (see [18]), the requested update and upgrade of FMVSS 111 is not likely to happen in the near future.

2.6 Exemption Procedures

In special cases (e.g. new technology which is not in line with the applicable regulations), it is possible to get a special authorization for a certain vehicle type. An example is the Volkswagen XL1 equipped with CMS instead of rearview mirrors and already produced in a small series (see [23]). In the following sections, an analysis of exemptions for new technologies (like e.g. CMS) will be presented. As such exemption procedures are country-specific, two examples are given by focusing on the European- and on the US-market.

2.6.1 Europe

Article 20: Exemptions for New Technologies or New Concepts

Article 20 of the European Directive 2007/46/EC for vehicle type approvals (see [4]) foresees the possibility of obtaining exemption approvals if a technology does not comply with one or more regulations that are required for type approval (see Fig. 16). The manufacturer will apply for an exemption in the EU member state of their choice. If the member state decides to support the application, it may grant a provisional approval, which is however restricted to the territory of the member state that granted the approval (the "grantor"). The other member states may then be approached by the grantor of the approval and be asked if they are also willing to accept the provisional approval.

In the next step, the grantor informs the European Commission and applies for an EC whole-vehicle type approval, including the new technology. The following information must be submitted to the EU-Commission:

- An explanation, why a certain requirement cannot be fulfilled;
- Supporting documents which show that the new technology ensures the same level of safety and environmental protection as the existing ones;

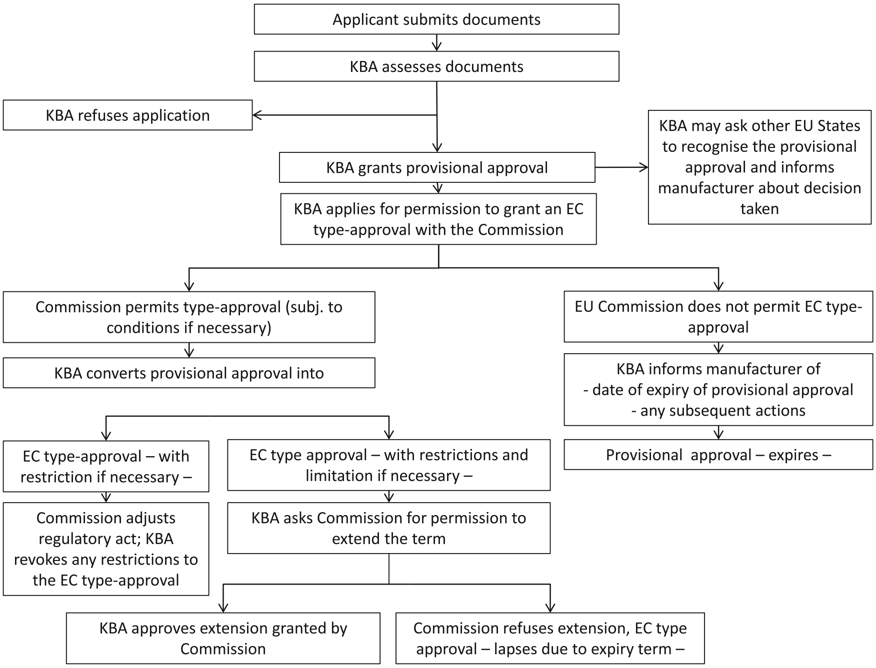


Fig. 16 Granting of EC type approval for new technologies or concepts, *source* [10], p. 5

- A description of the conducted tests and results which led to the conclusion that the same level of safety and environmental protection is ensured;

Based on the information provided, the EU Commission decides whether or not to authorize the application. The regulatory committee TCMV (Technical Committee Motor Vehicle), consisting of representatives of the EU Commission and the EU member states, assists the EU Commission in its decision-making. Due to a large number of stakeholders within the TCMV, the outcome of the decision-making is somewhat difficult to predict. The manufacturer cannot be certain whether the provisional approval will in the end become a regular vehicle type approval.

If TCMV grants the approval for exemption, it immediately has to start taking action for the amendment to the affected regulation in order to allow a certification without exemption procedures.

An approval for exemption is valid for a minimum of 36 months.

Article 23: National Small Series

According to Article 23 of the Directive 2007/46/EC, a manufacturer can apply for a national small series to EU member states. This procedure can also be a fallback solution if TCMV does not accept an exemption according to Article 20.

For small series, the authority can define alternative national requirements, when a specific regulation of the Directive 2007/46/EC cannot be complied with. In such

Table 4 Comparison of exemption procedures according to Directive 2007/46/EC, *source* Based on [23], p. 18

Article 20: New technologies	Article 23: National small series
– Ensure the same level of safety and environmental protection	– Ensure the same level of safety and environmental protection
– EU Commission is involved	– EU Commission is not involved
– Provisional type approval with national validity (approval by other EU member states is optional)	– National type approval without restrictions/limitations (approval by other EU member states is optional)
– After decision of the EU Commission: EC type approval is granted (may be subject to restrictions/limitations) or not granted (expiration of the provisional type approval)	– Number of vehicles is limited to only 100 per year per member state
– If EC type approval is granted: Update of the affected regulatory act	

a case, the member state requests the manufacturer submit the same information as mentioned above to explain that a same level of safety and environmental protection as the existing technologies is ensured.

A national type approval of small series is restricted within the territory of the member state that granted the approval, but the member state can individually send the small series approval to other member states and ask for their acceptance. For passenger cars (M1), a national small series is limited to only 100 vehicles per year per member state. It is therefore not a real option for mass production vehicles; it could be however, an option for a limited edition or for concept cars. Table 4 gives an overview of comparison between exemption procedures according to Article 20 and Article 23 of the Directive 2007/46/EC.

2.6.2 USA

According to 49 CFR part 555 (see [15]), a manufacturer can petition the National Highway Traffic Safety Administration (NHTSA) for an exemption from the Federal Motor Vehicle Safety Standards (FMVSS) if a new safety technology innovation does not comply with certain requirements of these safety regulations. The manufacturer must explain in detail from which standards and which particular requirements of the affected FMVSS regulation the exemption is requested. Further, they must prove that the new technology ensures the same level of safety as vehicles with technologies that comply with the regulation.

If NHTSA grants the exemption for a certain technology, the manufacturer is allowed to sell 2500 vehicles in the USA within any 12 months' period. Generally, the duration of the exemption is temporary for a maximum of two years. In special cases where two years are not economically reasonable, a maximum of three years are granted.

NHTSA is cautious about granting such exemptions and does not actively promote them, since due to the low number of vehicles it is difficult to establish statistically significant data that could prove the safety benefit of a technology.

2.7 Summary Current Regulatory Situation

Camera systems optionally installed as comfort systems in addition to the mandatory rearview mirrors (e.g. parking cameras) are not subject to current regulations, the exception being rearview cameras in the USA as of 2016.

Current regulations in countries that have not adopted UN Regulation No. 46 do not yet permit CMS to replace legally required mirrors on the driver and passenger sides, as well as inside mirrors for passenger cars. So far, CMS have only been permitted to replace the fender mirror, which may be necessary to comply with the Japanese close-proximity field of vision.

Some countries have established exemption procedures enabling new technologies not compliant with applicable vehicle regulations to be certified under defined conditions and restrictions. In such a case, evidence has to be shown to the authority that the new technology offers at least the equivalent safety level as existing technologies.

UN Regulation No. 46 is the first regulation worldwide to allow the replacement of the mandatory rearview mirrors by CMS. The details of the amended regulation are explained in Sect. 3.

3 Requirements for CMS in UN Regulation No. 46

3.1 General Aspects

This section outlines the general structure and the requirements of UN Regulation No. 46. It shall serve as a basis to understand the requirements that were developed for CMS.

Scope

UN Regulation No. 46 sets out provisions for devices for indirect vision and their installation in vehicles of the category M, N and L (with bodywork).

Devices for indirect vision are primarily mirrors and camera monitor systems.

Vehicle categories are defined in Resolution R.E.3. (see [25]): Category M includes vehicles, which are intended for the carriage of passengers—it includes passenger cars (M1) and buses (M2 and M3). Category N includes vehicles which are intended for the carriage of goods, meaning commercial vehicles—it is divided into the subcategories N1, N2 and N3, according to the maximum vehicle mass. Category L includes vehicles with less than four wheels (e.g. motorbikes) and some special lightweight and low speed/low power vehicles with four wheels.

Classes of Devices for Indirect Vision

In order to support drivers in driving and maneuvering their vehicles safely, a sufficient rearward field of vision is required. The characteristics of the vehicles (e.g. dimension, geometry, driver's seating position and relevant driving maneuvers) for and within each category can vary widely. When comparing passenger cars and commercial vehicles, it becomes obvious that devices for indirect vision have a significant importance in commercial vehicles because direct vision is rather limited and blind spots are larger. Also commercial vehicles have much larger dimensions which makes it more difficult to maneuver them safely. Different minimum fields of visions must therefore be attributed to each vehicle category. UN Regulation No. 46 includes seven different classes of devices for indirect vision (see Table 5) and regulates which of them are compulsory and which are optional for each of the different vehicle categories. For a typical passenger car (M1), one Class I and two Class III devices (each on driver and passenger side) are mandatory.

The amended UN Regulation No. 46 enables the use of camera monitor systems instead of mirrors for Class I, II, III and IV devices. Class V and VI CMS have already been possible since 2005 (see Sect. 1.7). Their requirements remained unchanged. The required mirrors for Class VII devices (only relevant for category L with body work) will be the only ones which cannot be replaced by CMS. This is partly due to the fact that the scope of ISO 16505 in its initial version from 2015 is limited to devices for vehicle categories M and N, and therefore does not include Class VII devices.

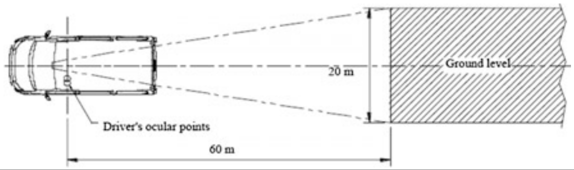
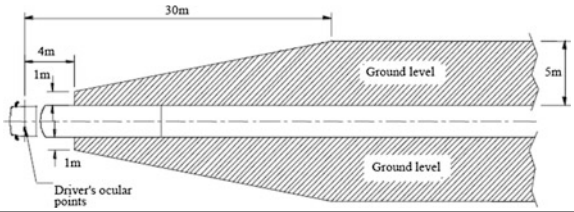
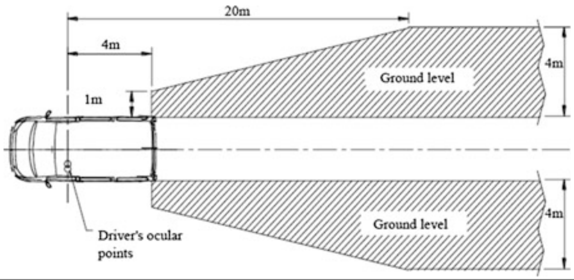
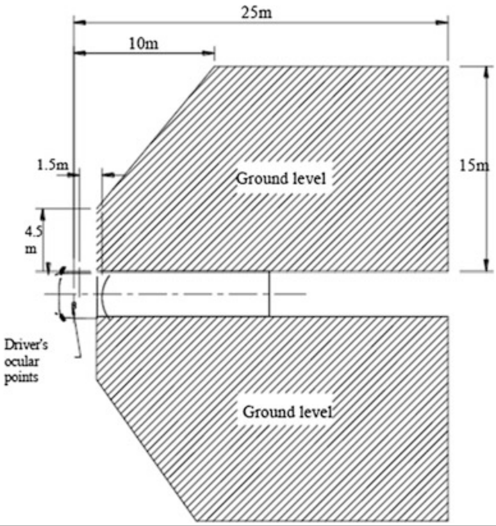
Requirements in UN Regulation No. 46

The requirements of UN Regulation No. 46 can be separated into two categories (see Fig. 17).

On the one hand there are technical requirements with regard to the component (device) itself. With the exception of the mechanical requirements (e.g. radius of external projections and impact absorption), the optical requirements and tests are individually designed for mirrors and camera monitor systems, because both systems have completely different characteristics that cannot be evaluated with the same tools and processes.

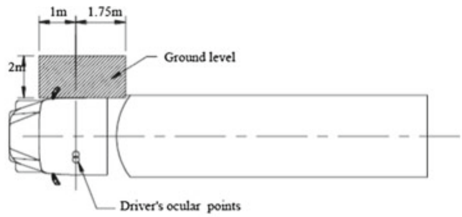
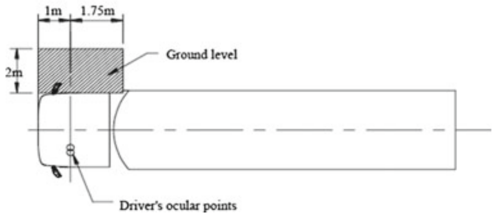
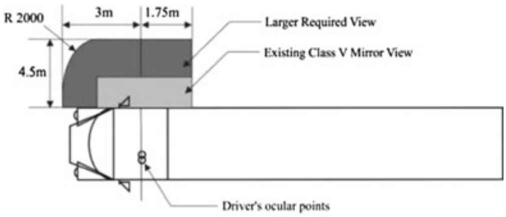
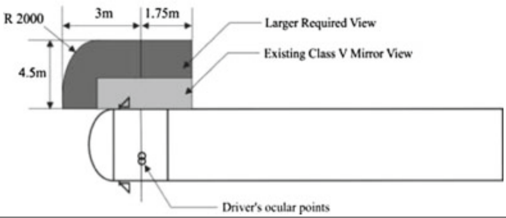
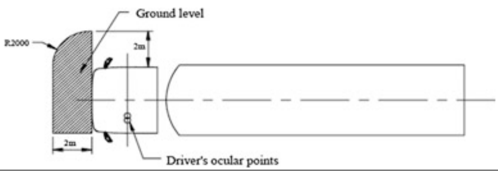
On the other hand, UN Regulation No. 46 includes requirements with regard to the installation of a device for indirect vision in a vehicle. The field of vision, which a device provides to the driver for a specific vehicle, mainly depends on its installation position. While the required fields of vision for each class are the same for mirrors and CMS, there are several other installation requirements that are specifically designed for the different technologies of mirrors and camera monitor systems.

Table 5 Overview of device classes and their required field of vision, *source* [32]

Class	Device	Field of vision
I	Rear view	
II	Main rear view	
III	Main rear view	
IV	Wide-angle view	
V	Close-proximity view	

(continued)

Table 5 (continued)

Class	Device	Field of vision
		(a) Devices which are situated less than 2.4 m above the ground   (b) Devices which are situated ≥ 2.4 m above the ground  
VI	Front view	

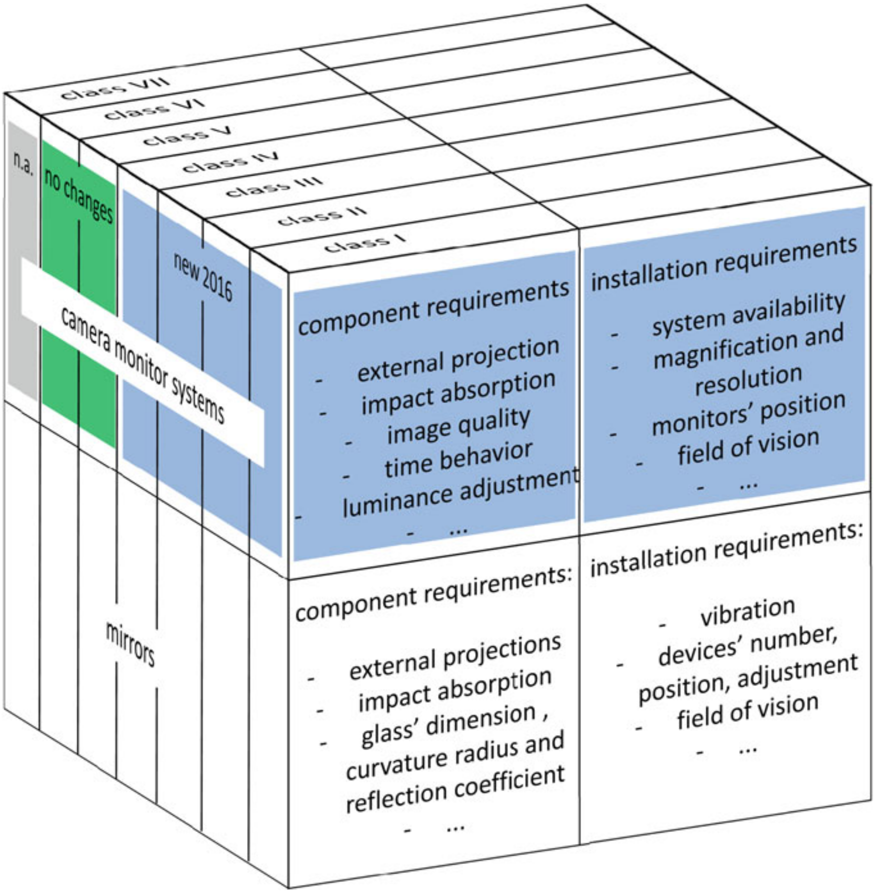


Fig. 17 Overview structure of UN Regulation No. 46

3.2 *Requirements for Class I–IV Camera Monitor Systems*

This section will give an overview of the new requirements for Class I–IV CMS as defined by the amended UN Regulation No. 46 (see [34]). It does not include each and every requirement and is not a substitution for the detailed study of the regulation necessary for the appropriate development and certification of a CMS. As explained in Sect. 1.7, the requirements and test methods of UN Regulation No. 46 are based on ISO 16505, but the final amendment draft included some modifications which were made by the Informal Group CMS. This section focuses on requirements which have a main impact on the implementation of a CMS into a vehicle. Furthermore, it gives an overview of different regulatory requirements and test methods in comparison to ISO 16505 including the impact on the system design that result from these modifications.

3.2.1 Mechanical Requirements

The mechanical requirements for CMS were basically carried over from mirrors to CMS. Accounting for the danger of head impact injury, the devices for indirect vision under UN Regulation No. 46 must have the following characteristics:

- Deflection occurs when hit by an obstacle (energy absorption) to reduce the impact,
- mounting may break away but without leaving sharp edges (radius > 2.5 mm),
- and additionally for camera monitor systems, the camera lens does not break.

These requirements are verified by a dynamic impact test with a pendulum that impacts the device. In order to pass the test, the pendulum must continue to swing after the impact within a minimum angle of 20° to the vertical. For this reason, outside mirrors for passenger cars (Class III) typically have a folding mechanism which ensures that the mirror housing deflects appropriately. CMS usually have a smaller camera holder design which does not foresee an installation space for a folding mechanism.

UN Regulation No. 46 includes some exemptions for the different mirror classes under which the impact test is not required. If, for instance, the installation height for devices of Class II–IV exceeds 2 m, the impact test is not required. In case of Class III mirrors for passenger cars, this exemption is practically irrelevant, but there are other configurations which are exempted from the pendulum test:

100 mm Exemption

- No impact test is required for devices that do not protrude "more than 100 mm measured beyond the circumscribing bodywork" (source [34]).
- The measurement is conducted as described in UN Regulation No. 26 (see [30]) (see Fig. 18). The protrusion is measured via the perpendicular relating to the circumscribing bodywork and the device must not protrude more than 100 mm into the respective direction.

Exemption for Integrated Devices with a Frontal Reflecting Area of Maximum 45°

- The device is integrated into the bodywork and its frontal reflecting area has an angle of less than 45° with the longitudinal median plane of the vehicle (see Fig. 19).
- In terms of complying with the regulation, a device is "integrated" when it is mounted flush with the body shell (e.g. no projecting bars, connecting links and eyes) and does not project beyond the bodywork in the corresponding section plane (see Fig. 20).

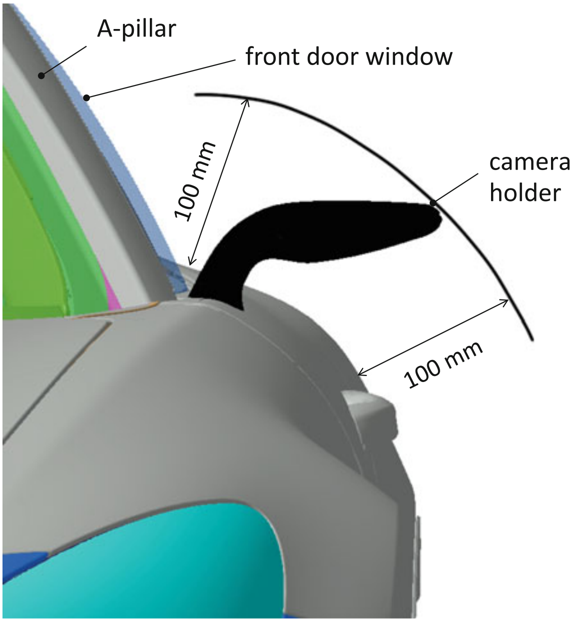


Fig. 18 How to measure the protrusion of parts (© Daimler AG)

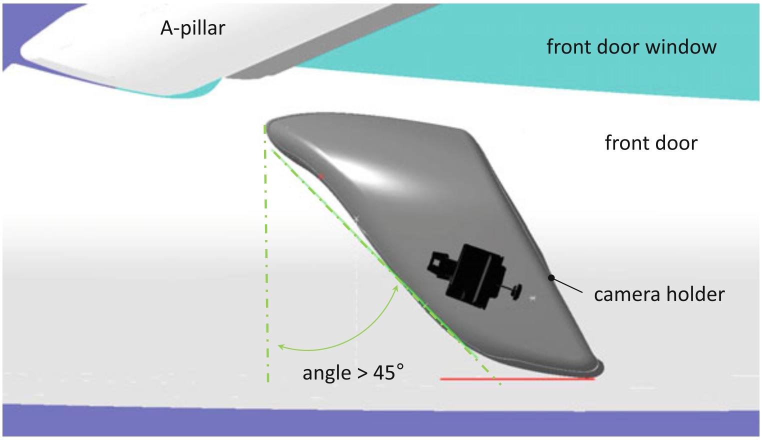


Fig. 19 Frontal reflecting area of max. 45° (© Daimler AG)

3.2.2 Functional Component Requirements

The functional component requirements focus on image quality. While the functional requirements for Class V and VI CMS mainly include luminance contrast and blooming, there are several new requirements for Class I–IV CMS due to different

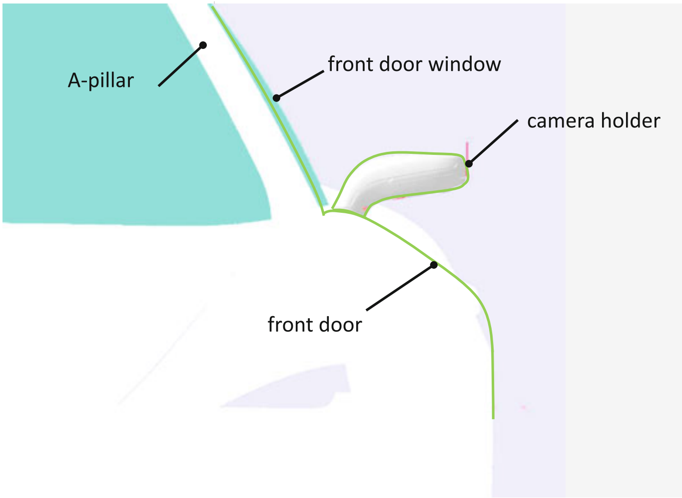


Fig. 20 Device integrated into the bodywork (© Daimler AG)

use-cases that have to be covered. For the definitions and symbols, the regulation refers to ISO 16505. The same applies to the test procedures unless otherwise specified in UN Regulation No. 46.

It should be noted that some component tests on image quality need vehicle-specific information: Standard and extended isotropy range depend on the monitors' design viewing direction, which in turn depends on the position of the monitor in relation to the driver's eye point. The new measurement procedure to determine the relevant luminance for the luminance contrast requirement depends also on the vehicle layout (see below).

Monitor Isotropy

The provisions regarding monitor isotropy are carried over from ISO 16505 without modifications.

Luminance and Contrast Rendering

The requirements in ISO 16505 are based on ISO 15008; UN Regulation No. 46 introduces several modifications.

While ISO 16505 requires a luminance contrast of 5:1 for night conditions, UN Regulation No. 46 sets out a more severe value of 10:1 (except for systems which include both technologies CMS and mirror) in order to address the needs of older drivers.

Furthermore, the test method for verifying the contrast under daylight conditions with diffuse ambient light has been changed. While ISO 16505 (as prescribed in ISO 15008) defines a light source of 1500 cd/m^2, UN Regulation No. 46 defines a range of from 1500–4200 cd/m^2 which takes into account some worst case scenarios. Meanwhile, UN Regulation No. 46 introduces a new procedure on how to define the applicable value within this range. This new procedure considers the fact that the relevant luminance, which monitors are exposed to, and its reflection to the driver's eyes are dependent on the whole vehicle layout (in particular the vehicle openings such as windows or sunroofs, including their dimensions and position).

Grey Scale Rendering

As ISO 16505 includes neither requirements nor test methods, both are newly introduced in UN Regulation No. 46.

Color Rendering

The provision regarding color rendering is carried over from ISO 16505 without modifications.

Artifacts: Smear

ISO limits smear to maximally 50 % of the maximum luminance value of the image. UN Regulation No. 46 defines a more severe value of maximal 10 % to be in line with the existing requirement for Class V and VI CMS.

Artifacts: Blooming and Lens Flare

Both ISO 16505 and UN Regulation No. 46 limit blooming and lens flare areas to maximally 25 % of the displayed image.

Artifacts: Point Light Sources

While ISO 16505 includes a qualitative requirement without test procedure, UN Regulation No. 46 introduces a severe quantitative requirement (two light sources with an intensity of 1750 cd and a distance of 1.3 m have to be distinguishable as two different light sources at a distance of 250 m) and a corresponding test procedure (point light source detection factor and point light source contrast factor).

Sharpness and Depth of Field

Both the requirements for sharpness and depth of field are carried over from ISO 16505 to UN Regulation No. 46 without modifications.

Geometric Distortion

While ISO 16505 includes a qualitative requirement, UN Regulation No. 46 defines a limit of 20 % relative to the recto-linear or pinhole projection for Class I, II and III CMS. Regarding the test method, UN Regulation No. 46 refers to Annex G.3. of ISO 16505.

Flicker

Both ISO 16505 and UN Regulation No. 46 require the monitor to be free from flicker for 90 % of the user population. While ISO 16505 does not include a test procedure, UN Regulation No. 46 introduces a test method by referring to the decision method that is described in Annex B of ISO 13406-2:2001.

Time Behavior: Frame Rate

The requirement is carried over from ISO 16505 to UN Regulation No. 46 without modifications.

Time Behavior: Image Formation Time

Both ISO 16505 and UN Regulation No. 46 require an image formation time of less than 55 ms at a temperature of 22 °C ± 5 °C. While ISO 16505 does not include a test procedure, UN Regulation No. 46 introduces one which is already included in ISO 9241-305:2008.

Time Behavior: System Latency

The requirement is carried over from ISO 16505 to UN Regulation No. 46 without modifications.

3.2.3 Installation Requirements

While the installation requirements for Class V and VI CMS focus mainly on the field of vision, there are several new requirements for Class I–IV CMS due to the different use-cases that are covered. For the definitions and symbols, the regulation refers to ISO 16505. The same applies to the test procedures unless otherwise specified in UN Regulation No. 46.

System Availability

Both UN Regulation R 46 and ISO 16505 require a warning to the driver if the system is not able to operate (e.g. CMS failure/malfunction). UN Regulation No. 46 requires that the field of vision is permanently visible to the driver as long as the ignition is on. A temporary switch-off therefore is not permitted, even in specific traffic situations where the vehicle does not move (e.g. stopped at a red light).

While ISO 16505 defines a switch-on time for a cold start for maximum of 7 s, UN Regulation No. 46 defines additional use-cases to ensure that the rearward field of vision is made available in a reasonable time frame, even if occupants of the vehicle do not leave immediately after switching off the engine (see Fig. 21). After each engine switch-off, the CMS shall continue to show the required field of vision for a minimum of 120 s (T1). Between T1 and a minimum of 300 s (T2 ≥ 420-T1), the entire field of vision must be made available within 1 s. UN Regulation No. 46 does not specify in detail how long the CMS must remain operational, if the engine is not switched on after the cold start. A reasonable approach in terms of complying with the regulation would be to start counting T1 as soon as the entire field of vision is available on the monitor after the cold start.

UN Regulation No. 46 foresees the possibility of allowing the manufacturer to deviate from the concept for activation and deactivation as shown in Fig. 21, provided that an alternative coherent safety concept can be shown and is accepted by the Technical Service.

Default View

The requirement is carried over from ISO 16505 to UN Regulation No. 46 without modifications.

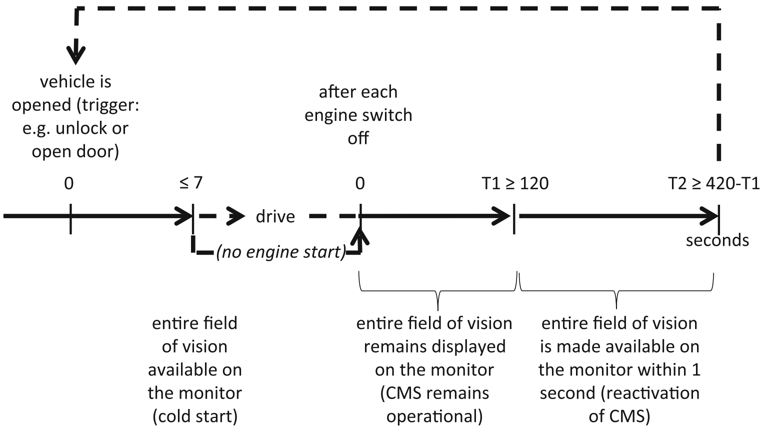


Fig. 21 System availability—activation and deactivation

Overlays

ISO 16505 permits overlays within the required field of vision as well, provided they are only temporary, transparent and related to driving. The two limiting aspects, "temporary" and "related to driving" are carried over to UN Regulation No. 46. It additionally requires that overlays are only allowed to display *rearward* driving-related information.

Regarding the size of the overlays, UN Regulation No. 46 stipulates more severe requirements which restrict the practical possibilities for system designs with overlays. While ISO 16505 does not regard transparent and temporary overlays as an obstruction, UN Regulation No. 46 considers any kind of overlay (regardless of their transparency) within the required field of vision as an obstruction. Therefore, UN Regulation No. 46 introduces limits for the allowed size of overlays:

The area size of each single overlay (e.g. icon, label, colored line etc.) is limited to a maximum of 2.5 % of the area which displays the minimum required field of vision.

In addition, the total area of all obstructions within the required field of vision is limited to a maximum of 15 % for devices of Class I and 10 % for all other device classes. It should be noted that these values already existed for mirrors. For Class I devices, they address obstructions due to backlight wipers, heating elements, etc. and for all other classes of devices obstructions due to bodywork, door handles etc. Therefore, not only overlays, but also all the other relevant obstructions fall under those limits. The amount of those other relevant obstructions mainly depends on the vehicle layout and positioning of the devices for indirect vision. Sports cars with wide rear fenders typically come close to the 10 % obstruction of the required field of vision in the outside mirror. In case of camera monitor systems that are supposed to have even shorter camera holders, the obstruction might even increase for bodyworks of the same type. Therefore, further obstructions in the form of reasonable overlays are unlikely to be included for some vehicle types. For the

Table 6 Comparison of the requirements for overlays within the required field of vision

ISO 16505	UN Regulation No. 46
Only transparent	Any overlays (regardless of their transparency) are considered as an obstruction
Only temporary	Only temporary
Only driving related visual information	Only rearward driving related visual information
Size unlimited	Maximal size of each overlay is 2.5 % of the area which displays the minimum required field of vision
	Overlays have to be included when calculating the obstructions within the required field of vision (obstructions are limited to a maximum of 15 % for devices of Class I and to a maximum of 10 % for all other device classes)

calculation of the overlays' total obstruction, the worst-case scenario is relevant, which means the biggest obstruction where one or more overlays are displayed at the same time. Table 6 gives an overview of comparison between the requirements for overlays of ISO 16505 and UN Regulation No. 46.

Magnification Factor

For CMS of Class I and III, the minimum magnification factor is carried over from ISO 16505 to UN Regulation No. 46 without modifications.

For CMS of Class II and IV, the minimum magnification factor in UN Regulation No. 46 is higher than that of the required values in chapter "Magnification and Resolution" of ISO 16505. The newly defined values are based on ISO's Annex 3, in which recommendations concerning commercial vehicles are formulated.

MTF (Resolution)

The requirement is carried over from ISO 16505 to UN Regulation No. 46 without modifications.

Magnification Aspect Ratio

The requirement is carried over from ISO 16505 to UN Regulation No. 46 without modifications.

Monitor Arrangement

The requirements of UN Regulation No. 46 are more restrictive than those of ISO 16505. While the location of the displayed field of visions in ISO 16505 is related to the monitor arrangement itself (left side field of vision to be displayed on the left side of the monitor arrangement and right side field of vision to be displayed on the right side of the monitor arrangement), UN Regulation No. 46 defines the allowed arrangement relative to the driver's ocular reference point (left side field of vision to be displayed left of the ocular reference point and right side field of vision to be displayed right of the ocular reference point). According to the UN Regulation No. 46 monitor arrangements similar to that of current outside mirrors may be

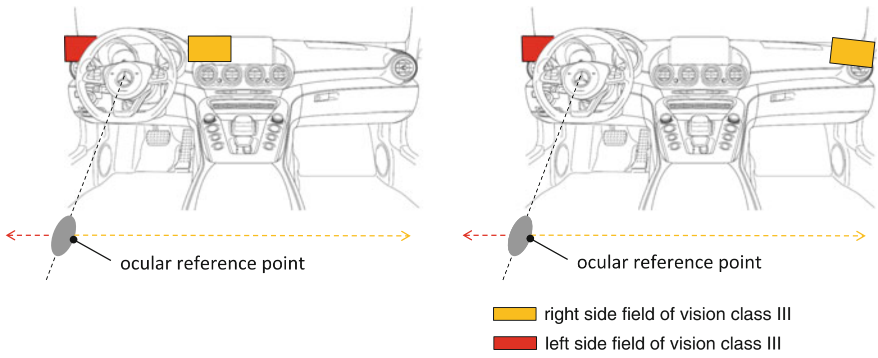


Fig. 22 Examples for permitted monitor arrangements of a Class III CMS (© Daimler AG)

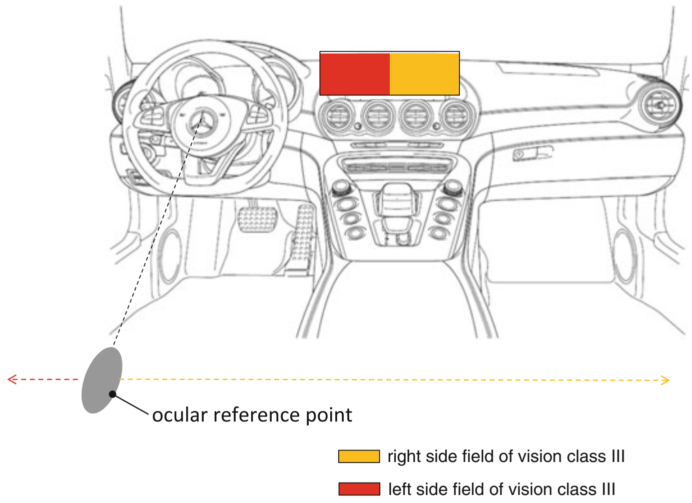


Fig. 23 Example for a monitor arrangement of Class III CMS which is not permitted by UN R-46 (© Daimler AG)

installed (see Fig. 22). Central monitor arrangements, which show both left and right side in a center position of the vehicle, are not permitted (see Fig. 23).

Decreasing Accommodation

The requirement is carried over from ISO 16505 to UN Regulation No. 46 without modifications.

Functional Safety

Because camera monitor systems are safety-relevant electronic systems, ISO 16505 generally requires their functional safety needs to be covered within the development process. ISO 16505 therefore refers to other existing standards that address functional safety.

UN Regulation No. 46 does not include this reference to other standards. Rather, it includes its own framework to support a methodology which is appropriate to avoid or control systematic and random failures of software and hardware. Therefore, it includes several criteria which are verified by the Technical Service for establishing an appropriate safety concept including documentation and verification. The key requirement of the safety concept is to ensure that the driver is informed (e.g. by a warning signal) in case of any electric/electronic malfunction of the CMS. A similar approach for functional safety was already successfully taken for UN Regulations No. 79 (Steering Equipment), No. 13 (heavy vehicle braking) and No. 13-H (braking of passenger cars) which include the so called Annex for complex electronic vehicle control systems to ensure a safe operation of such systems within the vehicle. Chapter "Functional Safety of Camera Monitor Systems" gives an in-depth analysis for functional safety topics of a CMS.

4 Other Relevant UN Regulations for CMS

Besides UN Regulation No. 46, there are some other UN Regulations which also affect the integration of a CMS into the vehicle. This section outlines other requirements that have to be taken into consideration within the development process to ensure product compliance with the relevant regulations.

4.1 UN Regulation No. 10 (Electromagnetic Compatibility)

UN Regulation No. 10 (see [26]) includes requirements regarding electromagnetic compatibility (EMC) i.e. emission of and immunity to electromagnetic radiation. As camera monitor systems are safety relevant electronic systems, their performance has to be resistant against magnetic or electrical fields. For that reason, UN Regulation No. 46 requires camera monitor systems to be in compliance with the technical requirements of Regulation No. 10.

4.2 UN Regulation No. 21 (Interior Fittings)

As explained in Sect. 3.2.1, UN Regulation No. 46 includes requirements to address, on the one hand, impacts of occupants on inside mounted devices for indirect vision (e.g. inside mirror Class I), and on the other hand, impacts of other road users (e.g. pedestrians) with outside mounted devices for indirect vision (e.g. outside mirror Class III). The impact requirement is verified with a dynamic pendulum test. UN Regulation No. 46 clarifies that monitors of a CMS are exempted from the pendulum test, as long as the vehicle has an approval according to UN Regulation No. 21.

UN Regulation No. 21 (see [29]) establishes requirements for avoiding injuries that result from an occupant's physical impact on the interior fittings. The regulation is limited to vehicles of category M1 (passenger cars). While the inside rearview mirror is exempted from UN Regulation No. 21 (rationale: energy dissipation is already covered by the pendulum test of UN Regulation No. 46), the CMS' monitors are within the scope of the regulation and therefore need to be considered. UN Regulation No. 21 includes several requirements (e.g. projections' radius of curvature, hardness of materials and energy dissipation of materials) that are related to different head impact zones. Therefore whether and which requirements and tests are applicable for CMS monitors' cannot be answered in general terms. This depends on the position of the monitors in the interior of the vehicle.

As a consequence, the integration of the monitors needs careful evaluation in order to ensure that the relevant head impact requirements of UN Regulation No. 21 can be fulfilled.

4.3 UN Regulation No. 48 (Installation of Lighting and Light-Signaling Devices)

UN Regulation No. 48 (see [35]) establishes requirements regarding the installation of lighting and light-signaling devices for vehicle of categories M and N, including their trailers (category O). The purpose of the regulation is to ensure, first, a minimum lighting of the road and, second, that a vehicle can be sufficiently recognized by other road users. For passenger cars (M1), the regulation requires e.g. the presence of two side direction indicator lamps. When installed on a vehicle, those lamps must comply with specific requirements on geometric visibility (angles) and photometry (illuminance). For some passenger cars, the side direction indicators are integrated into the housing of the outside mirrors (Class III). If mirrors are replaced by a CMS, the direction indicators need to be repositioned.

As a consequence, the repositioning of the direction indicators calls for careful evaluation in order to ensure that the relevant requirements of UN Regulation No. 48 are fulfilled.

4.4 UN Regulation No. 95 (Lateral Collision Protection)

UN Regulation No. 95 affects CMS such that after the impact test "no interior device or component shall have sharp projections or jagged edges which noticeably increase the risk of occupants' injury" (source [36], p. 12). This requirement is relevant for the integration of the monitors into the vehicle interior.

4.5 UN Regulation No. 125 (Forward Field of Vision)

UN Regulation No. 125 (see [27]) makes provisions on the 180° driver's field of vision to the front and is limited to vehicles which belong to category M1 (typical passenger cars). The purpose of the regulation is to ensure a sufficient direct field of vision for the driver.

Current Text of the Regulation

For the assessment of obstructions (e.g. caused by monitors) within the forward driver's field of vision, UN Regulation No. 125 defines

- three planes (to the front, left and right side) which start from eye point V2 and decline 4° to the horizontal
- a horizontal plane which starts from eye point V1

Between the 4° and the horizontal plane, no obstruction is allowed except for "A" pillars, vent window division bars, *rearview mirrors* and windscreen wipers (see Fig. 24).

As an alternative, UN Regulation No. 125 includes the option, that the conical projection of all obstructions (except for "A" pillars, vent window division bars, *rearview mirrors* and windscreen wipers) between the forward 4° plane and another plane, which declines 1° on a defined "area S" do not exceed more than 20 % of the total area S (see Fig. 25). This option only a ddresses obstructions that fall under the forward 4° plane. It cannot be applied for obstructions to the sides (e.g. upper edge of the doors).

The current text of the regulation only excludes rearview mirrors from the obstruction requirement. Other devices for indirect vision such as CMS fall under this requirement. Regarding CMS, the above-mentioned requirements would impact the integration of the monitors inside the vehicle and the camera-holder which is mounted to the vehicle exterior.

Impact on the Monitor Integration

Figure 26 shows an example of where the 4° plane is located in a typical sports car. Due to the low seating position and the relatively high upper edges of the front door, it becomes obvious that the limited available space between dashboard and

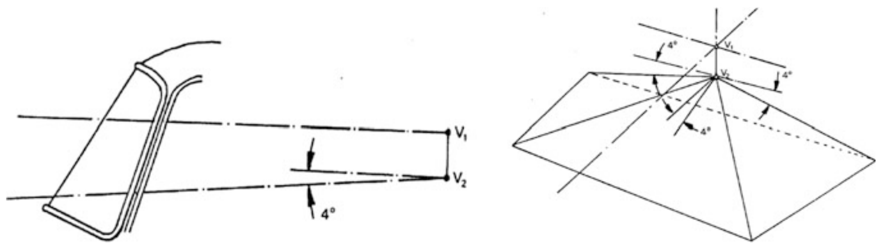


Fig. 24 Horizontal and 4° planes, *source* [27]

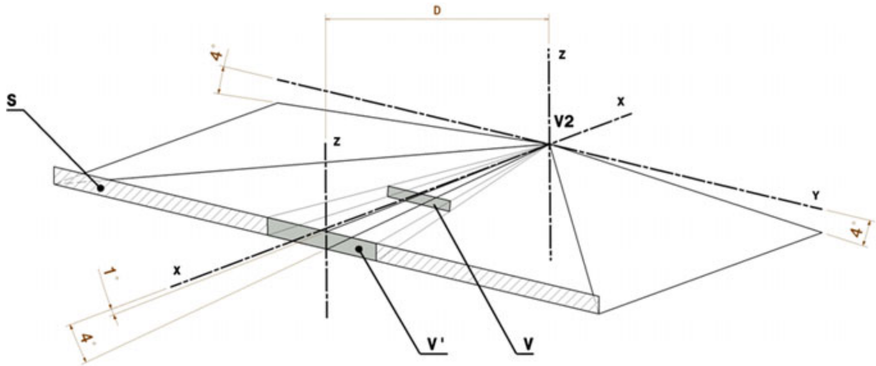


Fig. 25 Conical projection of obstructions between the 4° and 1° plane on area S, *source* [27]

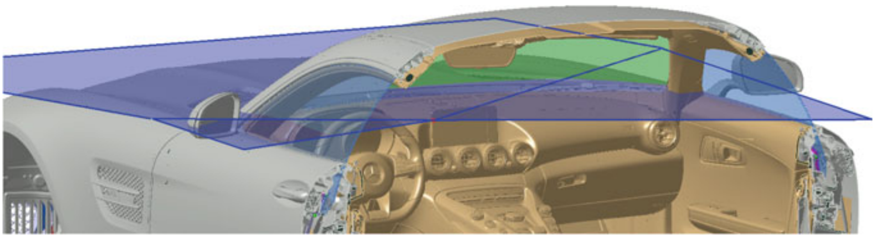


Fig. 26 4° planes at Mercedes-AMG GT (C190) (© Daimler AG)

those planes would not allow a reasonable integration of additional monitors for a CMS. The same applies for normal passenger cars.

Impact on the Camera Holder

As mentioned above, rearview mirrors are exempted from the obstruction assessment. Figure 27 shows an example where the outside mirrors (Class III) of a typical passenger car are located in relation to the 4° planes which decline to the side of the vehicle. In order to fulfill the required field of vision, rearview mirrors (and as well cameras) need to be mounted in a certain height above the ground level. For that reason, the outside mirrors on both driver and passenger side usually intrude above the 4° plane.

Amendment of Regulation No. 125 Towards CMS

In parallel to the amendment of UN Regulation No 46, a proposal for amendment of UN Regulation No. 125 was prepared by the Informal Group Camera Monitor Systems to address the above-mentioned conflicts with CMS (see [28]).

As a result, "cameras including their holders and housings which are mounted to the vehicle exterior" (source [28]) were added to the list of exemptions for the obstruction assessment.

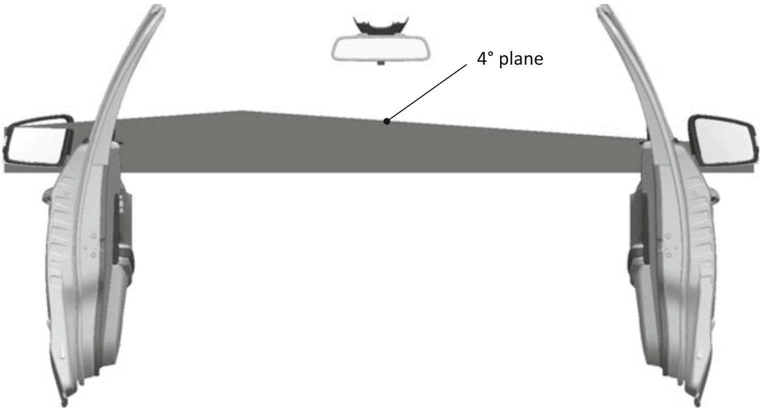


Fig. 27 Location of the outside mirrors relative to the 4° plane at Mercedes-Benz E-class (W212) (© Daimler AG)

While a certain installation height for the camera is needed to secure the required field of vision, for which reason an exemption is necessary, a monitor could have been installed in such a way that it fulfills the requirement for obstructions. Therefore, it was not the intention to add the monitors of a CMS to the list of exemptions in order to ensure that a CMS serves to improve the driver's forward field of vision. Compared to the current rearview mirrors, no improvement in terms of obstruction would have been ensured otherwise.

However, Fig. 26 shows that the required position of the monitors has a fundamental impact on the interior vehicle design, including the dashboard. To realize a display configuration, which complies with the obstruction requirements of Regulation No. 125, completely new interior design concepts are essential which can only be addressed at the beginning of a vehicle development cycle. For existing vehicles or vehicles in development however, such elementary design concepts cannot be addressed. In order to promote and support the installation of CMS on such vehicles, the proposal for amendment of UN Regulation No. 125 introduces a temporary exemption of the monitors (envisaged to expire in September 2021) under the following conditions:

- The vehicle is standard equipped with approved conventional rearview mirrors
- The approved mirrors are optionally replaced by a CMS
- The level of obstruction of the CMS does not exceed the level of obstruction of the mirrors
- The position of the monitors is as close as practical to the position of the rearview mirrors that are replaced.

Figure 28 shows a schematic example comparing the obstruction of a mirror and of a CMS. The monitor intrudes into the 4° plane to the front and to the side. However, the level of obstruction to the driver's forward field of vision is still much lower compared to the obstruction caused by a standard outside mirror. The figure

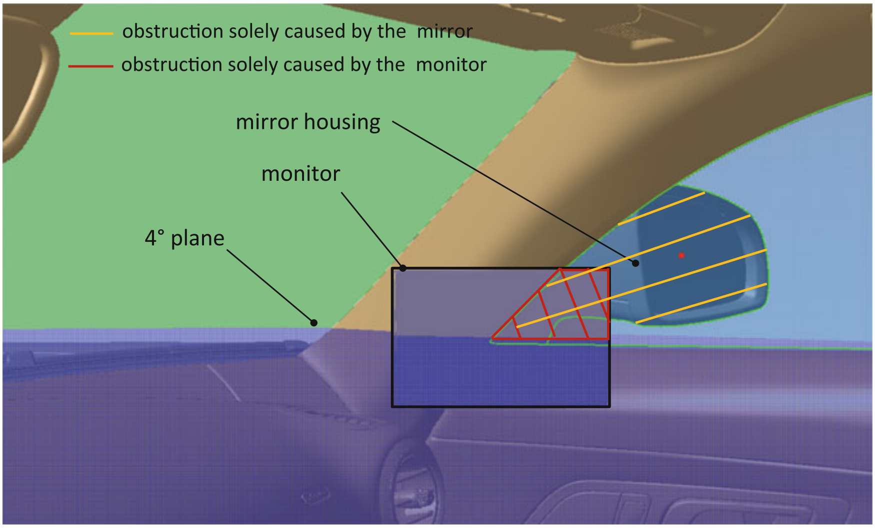


Fig. 28 Schematic comparison of the obstructions caused by a mirror and a monitor (© Daimler AG)

also illustrates the challenge to incorporate CMS monitors into current vehicle concepts in such a way that both ergonomic and design requirements are covered.

The proposal for the amendment of UN Regulation No. 125 (see [28]) was approved by GRSG in October 2015 and is likely to be adopted by WP.29 at its March session in 2016.

5 Conclusion and Future Activities

This final section will provide an overview of potential future activities on regulation and standardization of CMS.

Regulatory Activities in Other Regions

The amendment of UN Regulation No. 46 (see [34]) was a first step on the way to introducing CMS on a wider scale. Although UN Regulation 46 has been adopted by a significant number of countries (see Annex B), there are still some important markets such as China and the USA that are not covered by it. In a first step, the replacement of mirrors by CMS will most likely remain an option which can be offered in markets that accept the CMS technology. In other markets the same vehicle will have to be equipped with rearview mirrors.

As most of the vehicle manufacturers are keen to sell their products on a global scale, rearview mirrors will remain predominant unless the regulatory framework is harmonized. Otherwise, both technologies would have to be applied during the vehicle development phase, which is not economically feasible. Especially in Canada, China and in the USA, the issue must be addressed by regulatory activities. The regulatory development in other regions can benefit from the experience vehicle manufacturers will gain from the initial CMS applications, which are likely to be introduced in the UNECE market in the foreseeable future.

Follow-Up Activities on UN Regulation No. 46

The work on UN Regulation No. 46 is likely to be continued in the future. One possible working item could be Class VII CMS. Neither ISO 16505 nor UN Regulation No. 46 addresses Class VII CMS, which are for instance installed on motorbikes. Because the use-cases for motorbikes differ completely from those of other mirror classes, which among others, are installed on passenger cars or commercial vehicles, it seems reasonable that an ISO expert working group could conduct an initial analysis of individual safety needs for the application of Class VII devices. Based on the outcome, UN Regulation No. 46 can then be modified.

Another issue that calls for a reassessment is the arrangement of monitors as defined by UN Regulation No. 46. This regulation prohibits monitors that are installed in the center of the vehicle and enable the driver to capture both left and right side field of vision. Studies (e.g. [3]) conducted so far have focused on monitor positions similar to those of rearview mirrors to the right and left side of the driver's ocular reference point. For the reassessment, studies to analyze the suitability and acceptance of central monitor arrangements are needed. Because both technologies will be available on the market and car users may switch between cars equipped with either mirrors or CMS, it seems necessary to evaluate if drivers are able to adapt their driving behavior to different arrangements of monitors and mirrors when changing from one type of vehicle to another.

A third issue to be considered is the requirement for overlays. UN Regulation No. 46 defines any overlays as an obstruction in the driver's field of vision, regardless of their transparency. Due to the fact such obstructions are restricted and include obstructions of other parts like bodywork and door handles, the possibilities for the implementation of overlays are rather limited. It would therefore be necessary to evaluate to what extent an overlay with a certain transparency is a de facto negative obstruction in the field of vision, or rather a supporting element that helps the driver to maneuver and drive safely, without limiting the driver's perception in the required field of vision.

Other Relevant Regulations

Section 4 gave an overview of other UN Regulations which have an impact on the integration of CMS into a vehicle. In this context, the importance of UN Regulation No. 125 regarding the installation of monitors was highlighted. The amendment proposal (see [28]) excludes camera holders and cameras from the forward field of vision requirements. Furthermore, a temporary exemption was added to ease the integration of monitors, provided that their obstruction of the driver's forward field of vision does not exceed the obstruction due to standard rearview mirrors that are replaced. It should be noted, that other countries such as China (see [22]), India (see [2]) and South Korea (see [13]) have established their own national regulations for the forward field of vision of the driver. However, their requirements were carried-over in the past from UN Regulation No. 125. As a consequence, these regulations need to be updated according to the envisaged amendment of UN Regulation No. 125 (see [28]).

Future Revision of ISO 16505

The content of ISO standards needs to be reviewed after a maximum period of 5 years in order to clarify whether the standards still reflect the state of the art (see Sect. 1.2). As explained in Sect. 3, there are some differences between ISO 16505 and the amended UN Regulation No. 46. Some new requirements and test methods (e.g. point light sources) are included in UN Regulation No. 46. At the occasion of the next review of ISO 16505, the experts can have a closer look at the differences in order to work out potential harmonization aspects for both documents.

Annex A

See Table 7.

Table 7 Countries parties to the 1958 agreement*, *source* [38], pp. 39–40

ECE symbols	Contracting parties	Date of adhesion
E1	Germany[1]	28.01.1966
E2	France	20.06.1959
E3	Italy	26.04.1963
E4	Netherlands	29.08.1960
E5	Sweden	20.06.1959
E6	Belgium	05.09.1959
E7	Hungary	02.07.1960
E8	Czech Republic[3]	01.01.1993
E9	Spain	10.10.1961
E10	Serbia[10]	12.03.2001
E11	United Kingdom	16.03.1963
E12	Austria	11.05.1971
E13	Luxembourg	12.12.1971
E14	Switzerland	28.08.1973
E16	Norway	04.04.1975
E17	Finland	17.09.1976
E18	Denmark	20.12.1976
E19	Romania	21.02.1977
E20	Poland	13.03.1979
E21	Portugal	28.03.1980
E22	Russian Federation	17.02.1987
E23	Greece	05.12.1992

(continued)

Table 7 (continued)

ECE symbols	Contracting parties	Date of adhesion
E24	Ireland[9]	24.03.1998
E25	Croatia[5,15]	08.10.1991
E26	Slovenia[2]	25.06.1991
E27	Slovakia[4]	01.01.1993
E28	Belarus	02.07.1995
E29	Estonia	01.05.1995
E31	Bosnia and Herzegovina[6]	06.03.1992
E32	Latvia	18.01.1999
E34	Bulgaria	21.01.2000
E35	Kazakhstan	08.01.2011
E36	Lithuania	29.03.2002
E37	Turkey	27.02.1996
E39	Azerbaijan	14.06.2002
E40	The Former Yugoslav Republic of Macedonia[7]	17.11.1991
E42	European Union[8]	24.03.1998
E43	Japan	24.11.1998
E45	Australia	25.04.2000
E46	Ukraine	30.06.2000
E47	South Africa	17.06.2001
E48	New Zealand	26.01.2002
E49	Cyprus[11]	01.05.2004
E50	Malta[11]	01.05.2004
E51	Republic of Korea	31.12.2004
E52	Malaysia[12]	04.04.2006
E53	Thailand[13]	01.05.2006
E54	Albania	05.11.2011
E56	Montenegro[14]	03.06.2006
E58	Tunisia	01.01.2008
E60	Georgia	25.05.2015
E62	Egypt	03.02.2013

*A daily updated list of the Contracting Parties to the Agreement is available at: http://www.unece.org/trans/conventn/agreem_cp.html#18

[1]Effective 3 October 1990, the German Democratic Republic acceded to the Federal Republic of Germany

[2]Succession to Yugoslavia, Depositary notification C.N.439.1992.TREATIES-53 of 18 Mar 1993

[3]Succession to Czechoslovakia, Depositary notification C.N.229.1993.TREATIES of 14 December 1993

[4]Succession to Czechoslovakia, Depositary notification C.N.184.1993.TREATIES, received on 20 July 1994

[5]Succession to Yugoslavia, Depositary notification C.N.66.1994.TREATIES-10 of 31 May 1994

[6]Succession to Yugoslavia, Depositary notification C.N.35.1994.TREATIES of 2 May 1994

[7]Succession to Yugoslavia, Depositary notification C.N.142.1998.TREATIES-33 dated 4 May 1998

[8]Approvals are granted by its Member States using their respective ECE symbol

[9]Not Contracting Party to the Agreement, but by virtue of accession to the Agreement by the European Union on 24 March 1998, Irland applies
the same UN Regulations than the European Union
[10]Succession to Yugoslavia, Depositary notification C.N.276.2001.TREATIES-3 dated 2 April 2001
[11]Not Contracting Parties to the Agreement, but by virtue of accession to the Agreement by the European Union on 1 May 2004, Cyprus and
Malta apply the same UN Regulations than the European Union
[12]Not bound by Article 10 of the Agreement
[13]Not bound by any of the UN Regulations, nor by Article 10 of the Agreement
[14]Succession to Yugoslavia, Depositary Notification C.N.1346.2006.TREATIES-3 dated 1 Mar 2007
[15]By virtue of accession to the European Union on 1 July 2013, Croatia applies the same UN Regulations than the European Union

Annex B

See Table 8.**Table 8** Signatories of UN Regulation No. 46, *source* [38], p. 89

ECE symbols	Contracting parties	Date of application
E1	Germany	20.04.1986
E2	France	01.09.1981
E3	Italy	01.09.1981
E4	Netherlands	04.12.1987
E5	Sweden	24.09.1982
E6	Belgium	16.10.1982
E7	Hungary	26.03.1984
E8	Czech Republic	18.09.1982
E9	Spain	24.03.1989
E10	Serbia	18.05.2008
E11	United Kingdom	27.04.1990
E12	Austria	23.07.1990
E13	Luxembourg	01.10.1983
E14	Switzerland	not yet signed
E16	Norway	24.05.1993
E17	Finland	10.08.1982
E18	Denmark[1]	24.03.1998
E19	Romania	03.02.1984
E20	Poland	03.06.1990
E21	Portugal[1]	24.03.1998
E22	Russian Federation	06.03.1988
E23	Greece	03.12.1995

(continued)

Table 7 (continued)

ECE symbols	Contracting parties	Date of application
E24	Ireland[1]	24.03.1998
E25	Croatia	03.04.2001
E26	Slovenia	01.10.1994
E27	Slovakia	18.09.1982
E28	Belarus	02.07.1995
E29	Estonia	25.07.1999
E31	Bosnia and Herzegovina	not yet signed
E32	Latvia	18.01.1999
E34	Bulgaria[4]	01.01.2007
E35	Kazakhstan	not yet signed
E36	Lithuania	29.03.2002
E37	Turkey	07.07.2000
E39	Azerbaijan	not yet signed
E40	The Former Yugoslav Republic of Macedonia[7]	19.08.2002
E42	European Union[2]	24.03.1998
E43	Japan	not yet signed
E45	Australia	not yet signed
E46	Ukraine	08.10.2002
E47	South Africa	17.06.2001
E48	New Zealand	19.03.2002
E49	Cyprus[3]	01.05.2004
E50	Malta[3]	01.05.2004
E51	Republic of Korea	not yet signed
E52	Malaysia	04.04.2006
E53	Thailand	not yet signed
E54	Albania	05.11.2011
E56	Montenegro	not yet signed
E58	Tunisia	not yet signed
E60	Georgia	25.05.2015
E62	Egypt	03.02.2013

[1]By virtue of accession to the Agreement by the European Union
[2]Approvals are granted by its Member States using their respective ECE symbol
[3]By virtue of accession to the European Union on 1 May 2004
[4]By virtue of accession to the European Union on 1 January 2007

References

1. Alliance of Automobile Manufacturers and Tesla Motors, Inc.: Petition for Rulemaking to Update and Upgrade Federal Motor Vehicle Safety Standard No. 111, "Rearview mirrors" to allow Camera-based Systems as a Compliance Option, March 28, 2014 http://www.autoalliance.org/download.cfm?downloadfile=3863B6A0-BEAF-11E3-AB0B000C296BA163&typename=dmFile&fieldname=filename (2014). Accessed 14 Oct 2015
2. Central Motor Vehicle Rules: Automotive Industry Standard AIS-021/2014, Field of Vision of Motor Vehicle Drivers for M1 Category, December 2004, Last amendment: Amendment No. 2 of 2015-04-10, Ministry of Shipping, Road Transport & Highways, Copyright© 2015 by: The Automotive Research Association of India (2015)
3. Federal Highway Research Institute (BASt): Schmidt, Eike Andreas; Hoffmann, Heike; Krautscheid, Rainer; Bierbach, Maxim; Frey, Alexander; Gail, Jost; Lotz-Keens, Christiane Final Report, Camera Monitor Systems as a Replacement for Exterior Mirrors in Cars and Trucks, 22.01.2015 https://www2.unece.org/wiki/download/attachments/25264588/IGCMS-II-07-17%20BASt%20Final_Report_CMS.PDF (2015). Accessed 14 Oct 2015
4. European Parliament and Council: Directive 2007/46/EC of 5 September 2007 establishing a framework for the approval of motor vehicles and their trailers, and of systems, components and separate technical units intended for such vehicles, last amendment (EU) No. 195/2013 of 7 March 2013 http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:02007L0046-20150224&qid=1437644384084&from=EN (2007). Accessed 14 Oct 2015
5. International Organization for Standardization: Standards. What is a standard? Copyright remains with ISO. http://www.iso.org/iso/home/standards.htm (2015). Accessed 14 Oct 2015
6. International Organization for Standardization: Who develops ISO standards? Copyright remains with ISO. http://www.iso.org/iso/home/standards_development/who-develops-iso-standards.htm (2015). Accessed. 14 Oct 2015
7. International Organization of Standardization: My ISO job. What delegates and experts need to know. ISO, 2015. Copyright remains with ISO. http://www.iso.org/iso/my_iso_job.pdf ISBN 978-92-67-10632-8. Accessed 14 Oct 2015
8. International Organization of Standardization: International harmonized stage codes. Copyright remains with ISO. http://www.iso.org/iso/stage_codes.pdf (2015). Accessed 14 Oct 2015
9. International Organization of Standardization: International Standard ISO 16505, First edition 2015-05-01, Road vehicles—Ergonomic and performance aspects of Camera Monitor Systems—Requirements and test procedures. Copyright remains with ISO. Reference number ISO 16505:2015(E) (2015)
10. Kraftfahrtbundesamt: Information Sheet on Approvals for "New Technologies or Concepts" (MTK), Issue July 2011 http://www.kba.de/EN/Fahrzeugtechnik_en/Zum_Herunterladen_en/zum_Herunterladen_node_en.html (2011). Accessed 14 Oct 2015
11. Japan Ministry of Land, Infrastructure and Transport: MLIT Announcement 619 of July 15, 2002 (TS Mirrors for confirming the immediate front and left side), Last amended by Announcement No. 872 of August 16, 2005.), Copyright © 2006 by: Japan Automobile Standards Internationalization Center (http://www.jasic.org/)
12. Japan Ministry of Land, Infrastructure and Transport: Safety Regulations for Road Vehicles, Safety Regulation Art. 44 (Ord. 67/1951), Last amendment Ordinance No. 85 of 2008, Article 44 (Rear-view Mirrors, etc.), Copyright © 2010 by: Japan Automobile Standards Internationalization Center (http://www.jasic.org/)
13. Ministry of Land, Transport and Maritime Affairs: Regulations for Motor Vehicle Safety Standards, KMVSS Art. 94 (MLTM Not. 515/2012), Field of view: Direct vision, English version translated from Korean original, last amendment: Ministry of Land, Transport and Maritime Affairs Notification No. 515 (2012-08-31)

14. Ministry of Land, Infrastructure and Transport: Regulations for Motor Vehicle Safety Standards, KMVSS Art. 50 (MOLIT Ord. 164/2014), Field of view: Indirect vision, English version translated from Korean original, last amendment: MOLIT Ordinance No. 164 (2014-12-31)
15. Department of Transportation, National Highway Traffic Safety Administration: Part 555—Temporary Exemption From Motor Vehicle Safety And Bumper Standards http://www.gpo.gov/fdsys/pkg/CFR-2011-title49-vol6/pdf/CFR-2011-title49-vol6-part555.pdf (1973). Accessed 14 Oct 2015
16. Department of Transportation, National Highway Traffic Safety Administration: §571.111 Standard No. 111; Rearview mirrors http://www.gpo.gov/fdsys/pkg/CFR-2011-title49-vol6/pdf/CFR-2011-title49-vol6-sec571-111.pdf (2011). Accessed 14 Oct 2015
17. Department of Transportation, National Highway Traffic Safety Administration: 49 CFR Part 571, [Docket No. NHTSA-2010-0162], Federal Motor Vehicle Standards, Rear Visibility, Final Rule from April 7, 2014, Federal Register http://www.gpo.gov/fdsys/pkg/FR-2014-04-07/pdf/2014-07469.pdf (2014). Accessed 14 Oct 2015
18. National Highway Traffic Safety Administration: Overview of NHTSA Priority Plan for Vehicle Safety and Fuel Economy, 2015-2017, June 2015 http://www.nhtsa.gov/staticfiles/nvs/pdf/NVS_priority-plan-June2015_final.pdf (2015). Accessed 14 Oct 2015
19. Organisation Internationale des Constructeurs d'Automobile: OICA Members http://www.oica.net/category/about-us/members/ (2015). Accessed 14 Oct 2015
20. Organisation Internationale des Constructeurs d'Automobile: Technical Committee http://www.oica.net/category/about-us/committees/technical-committee/ (2015). Accessed 14 Oct 2015
21. People's Republic of China: National Standard of the People's Republic of China, GB 15084-2013, Motor vehicles-devices for indirect vision—requirements of performance and installation, issued on 2013-09-18, effective from 2014-07-01, Promulgated by State General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China And Standardization Administration of China, ACEA Translation from Chinese original, For Reference Purposes Only (2013)
22. People's Republic of China: National Standard of the People's Republic of China, GB 11562-2014, Motor vehicles-forward visibility for drivers-requirements and measurement methods, issued on 2014-12-31, effective from 2015-07-01, Promulgated by State General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China And Standardization Administration of the People's Republic of China, ACEA Translation from Chinese original, For Reference Purposes Only (2014)
23. TÜV-NORD Mobilität GmbH & Co KG: Burak Kaya, Typgenehmigungsverfahren im Fall neuer Techniken am Beispiel des Volkswagen XL1, 6. Sachverständigentag, 17 March 2015 http://www.vdtuev.de/sachverstaendigentag/praesentationen/dok_view?oid=515508&vater=515003 (2015). Accessed 14 Oct 2015
24. United Nations Department of Public Information: The United Nations System, DPI/2470—10-00133—April 2011 http://www.un.org/depts/dhl/deplib/promo-materials/UNsystemChart-lt-clr.pdf (2011). Accessed 14 Oct 2015
25. United Nations: ECE/TRANS/WP.29/78/Rev.3, 23 January 2014, Consolidated Resolution on the Construction of Vehicles (R.E.3.), Revision http://www.unece.org/fileadmin/DAM/trans/main/wp29/wp29resolutions/ECE-TRANS-WP29-78-r3e.pdf (2014). Accessed 14 Oct 2015
26. United Nations: E/ECE/324/Add.9/Rev.5 - E/ECE/TRANS/505/Add.9/Rev.5, 16 October 2014, Addendum 9: Regulation No. 10, Revision 5, Uniform provisions concerning the approval of motor vehicles with regard to electromagnetic compatibility http://www.unece.org/fileadmin/DAM/trans/main/wp29/wp29regs/updates/R010r5e.pdf (2014). Accessed 14 Oct 2015
27. United Nations: E/ECE/324/Rev.1/Add.124/Rev.2, 19 August 2013, Addendum 124: Regulation No. 125, Revision 2, Uniform provisions concerning the approval of motor vehicles with regard to the forward field of vision of the motor vehicle driver http://www.unece.org/fileadmin/DAM/trans/main/wp29/wp29regs/2013/R125r2e.pdf (2013). Accessed 14 Oct 2015

28. United Nations: GRSG-109-23 Rev.1, Proposal for the 01 series of amendments to Regulation No. 125 (Forward field of vision of drivers) http://www.unece.org/fileadmin/DAM/trans/doc/2015/wp29grsg/GRSG-109-23r1e.pdf (2015). Accessed 14 Oct 2015
29. United Nations: E/ECE/324/Rev.1/Add.20/Rev.2/Amend.2, E/ECE/TRANS/505/Rev.1/Add.20/Rev.2/Amend.2, 25 March 2003, Addendum 20: Regulation No. 21, Revision 2—Amendment 2, Uniform provisions concerning the approval of motor vehicles with regard to their interior fittings http://www.unece.org/fileadmin/DAM/trans/main/wp29/wp29regs/r021r2a2e.pdf (2003). Accessed 14 Oct 2015
30. United Nations: E/ECE/324/Rev.1/Add.25/Rev.1, 19 August 2013, Addendum 25: Regulation No. 26, Revision 1, Uniform provisions concerning the approval of vehicles with regard to their external projections http://www.unece.org/fileadmin/DAM/trans/main/wp29/wp29regs/2013/R026r1e.pdf (2013). Accessed 14 Oct 2015
31. United Nations: GRSG-98-16, Regulation No. 46 (Devices for indirect vision), Report of the informal group on camera monitor systems from 29 April 2010 http://www.unece.org/fileadmin/DAM/trans/doc/2010/wp29grsg/GRSG-98-16e.pdf (2010). Accessed 14 Oct 2015
32. United Nations: E/ECE/324/Rev.1/Add.45/Rev.5, 19 August 2013, Addendum 45: Regulation No. 46, Revision 5, Uniform provisions concerning the approval of devices for indirect vision and of motor vehicles with regard to the installation of these devices http://www.unece.org/fileadmin/DAM/trans/main/wp29/wp29regs/2013/R046r5e.pdf (2013). Accessed 14 Oct 2015
33. United Nations: GRSG-108-46-Rev.2, Proposal for a draft Supplement to the 04 series of amendments to Regulation No. 46 (Devices for indirect vision) http://www.unece.org/fileadmin/DAM/trans/doc/2015/wp29grsg/GRSG-108-46r2e.pdf (2015). Accessed 14 Oct 2015
34. United Nations: ECE/TRANS/WP.29/2015/84, 27 August 2015, Proposal for Supplement 2 to the 04 series of amendments to Regulation No. 46 (Devices for indirect vision) http://www.unece.org/fileadmin/DAM/trans/doc/2015/wp29/ECE-TRANS-WP29-2015-084e.pdf (2015). Accessed 14 Oct 2015
35. United Nations: E/ECE/324/Rev.1/Add.47/Rev.12 - E/ECE/TRANS/505/Rev.1/Add.47/Rev.12, 16 October 2014, Addendum 47: Regulation No. 48, Revision 12, Uniform provisions concerning the approval of vehicles with regard to the installation of lighting and light-signalling devices http://www.unece.org/fileadmin/DAM/trans/main/wp29/wp29regs/updates/R048r12e.pdf (2014). Accessed 14 Oct 2015
36. United Nations: E/ECE/324/Add.94/Rev.2 - E/ECE/TRANS/505/Rev.1/Add.94/Rev.2, 13 February 2014, Addendum 94: Regulation No. 95, Revision 2, Uniform provisions concerning the approval of motor vehicles with regard to the protection of the occupants in the event of a lateral collision http://www.unece.org/fileadmin/DAM/trans/main/wp29/wp29regs/updates/R095r2e.pdf (2014). Accessed 14 Oct 2015
37. United Nations: Agreement concerning the adoption of uniform technical prescriptions for wheeled vehicles, equipment and parts which can be fitted and/or be used on wheeled vehicles and the conditions for reciprocal recognition of approvals granted on the basis of these prescriptions. Revision 2. Including the amendments entered into force on 16 October 1995. http://www.unece.org/fileadmin/DAM/trans/main/wp29/wp29regs/505ep29.pdf (1958). Accessed 14 Oct 2015
38. United Nations: ECE/TRANS/WP.29/343/Rev.23, 2 September 2015, Status of the 1958 Agreement, of the annexed UN Regulations and of amendments thereto, Revision 23 http://www.unece.org/fileadmin/DAM/trans/main/wp29/wp29regs/updates/08-ECE-TRANS-WP.29-343-Rev.23.pdf (2015). Accessed 14 Oct 2015
39. United Nations: Mission http://www.unece.org/mission.html (2015). Accessed 14 Oct 2015
40. Verband der Automobilindustrie: Normenausschuss Automobiltechnik. Jahresbericht 2015 https://www.vda.de/de/services/Publikationen/naautomobil-jahresbericht-2015.html (2015). Accessed 14 Oct 2015

41. Verband der Automobilindustrie: Grundsätze der Normung https://www.vda.de/de/verband/fachabteilungen/fachabteilung-normung/grundsaetze-der-normung.html (2015). Accessed 14 Oct 2015
42. United Nations: World Forum For Harmonization of Vehicle Regulations (WP.29). How It Works. How To Join It. Third edition. New York and Geneva, 2012. http://www.unece.org/fileadmin/DAM/trans/main/wp29/wp29wgs/wp29gen/wp29pub/WP29_Blue_Book_2012_ENG.pdf (2012). Accessed 14 Oct 2015

Resolution and Sharpness Requirements for CMS

Eiji Oba

Abstract This chapter provides an introduction to the resolution and sharpness requirements for a camera monitor system = CMS (Camera Monitor System) to replace the traditional optical rear-view mirrors. It provides information on how the requirements are translated from what a driver could observe as the minimum angular detail through a mirror to what a driver can expect to observe on a monitor of a CMS. The minimum requirements are derived from the visual acuity of a licensed driver and the magnification of type approved mirrors. Further, this chapter provides guidance on measurement procedure and practical evaluation of a CMS. This chapter is aimed to help the reader to understand the minimum requirements determined by regulation and understand the respective measurement procedure of resolution and sharpness. It clarifies the difference between the units used in definition of the resolution requirements of the CMS and the resolution units used in a standard resolution measurement according to ISO 12233:2014. The evaluation procedure describes a method to verify how many distinguishable line widths are observable within the width in the orientation of interest. It provides a guide to evaluate the system directly on a partially captured image on the CMS monitor, independent of the output image aspect ratio or magnification aspect ratio.

Keywords Resolution · Visual acuity · MTF (Modulation Transfer Function) · SFR (Spatial Frequency Response) · Line widths per picture height · Resolution chart · Image height · Nominal spatial frequency · Normalized spatial frequency

1 Introduction

The spatial resolution of the CMS (Camera Monitor System) is one of the key attribute for performance of a camera monitor system among others. This chapter provides an introduction to the resolution (MTF) requirement given in ISO

E. Oba (✉)
Atsugi Technology Center, Sony Corporation, Atsugi, Japan
e-mail: eiji.oba@jp.sony.com

A. Terzis (ed.), *Handbook of Camera Monitor Systems*,
Augmented Vision and Reality 5, DOI 10.1007/978-3-319-29611-1_3

16505:2015 [1], including explanation on how the resolution and sharpness requirement were elaborated within this International Standard. The resolution (MTF) described in ISO 16505:2015 is not the addressable number of photo-elements but a parameter representing the integral performance of the CMS.

When a driver observes a scene through a convex mirror, the resulting minimum angular acuity of the viewing scene is degraded by the magnification factor of the convex mirror. Therefore, the effective magnification factor that a driver observes through a traditional mirror was studied in order to determine the requirements for a CMS, as replacement for traditional mirrors.

Magnification varies with the observation condition through the mirror, depending on the mirror curvature radius, on the viewing angle, on the distance from the mirror surface to the eye point location and on the distance to the observed scene. To introduce requirement in analogy to traditional mirrors, additional information of type approved vehicle by the time of ISO 16505 development in 2013 were collected to determine the typical magnification a driver observes using a traditional mirror. The collected information was analysed to determine the minimum and maximum viewing angle and create the range of the viewing angles, namely the window a driver would use to observe a backward scene through a mirror. Objects at far distance look small and are difficult to recognize while objects at closer distance look large and are easy to recognize. Hence, a higher priority was given to observe objects at far distance with magnification plots represented with distance at infinity. The “mirror average magnification factors” given under Sects. 3.2.23 and 3.2.24 of ISO 16505:2015 are such values derived from calculated magnification plot curves representing typical magnification at distance. Details of the magnification calculation procedure are given in the Annex B.3 of ISO 16505:2015.

The human eye’s visual acuity is traditionally checked using Landolt C chart and this chart is largely adopted as part of driver license test. The requirements described under ISO 16505:2015 target to achieve comparable scene distinguishability to what a driver with the minimal admitted visual acuity could observe using a traditional mirror; and secure that a CMS would provide scene distinguishability comparable to what a driver with minimum visual acuity could expect to see with a traditional convex mirror. However, due to the spatial discrete sampling characteristics of imaging technologies used in today’s imaging devices, the Landolt C chart does fit well for quantitative evaluation of CMS system. As alternative, Modulation Transfer Function (MTF) described in ISO 12233:2014 [2] was adopted for evaluating the requirement of CMSs. This standard make use of the full image in the measurement procedure, whereas the requirements for a CMS in ISO 16505:2015 are defined using a partial image. More precisely, the resolution requirement in ISO 16505:2015 is defined in a partial square image displayed on the CMS, and it is not defined in the actual full monitor image size. This difference in the definition of image size has a large impact on the nominal resolution value when measured in [LW/PH] units. Details of inconvenience will be explained in Sect. 3. It is suggested to understand the definition of the CMS requirements firstly without going into detail of the resolution measurement following exactly ISO

12233:2014 standard, which uses the full image size as a reference of the [LW/PH] units.

It is one aim of this handbook to help readers to understand the CMS requirements, and then provide information on how to evaluate a CMS and still take advantages of the well-established ISO 12233:2014 appropriately. A table is provided in Sect. 6 for comparison of differences in ISO 16505:2015 and ISO 12233:2014.

ISO 16505:2015 handles the resolution requirement based on the idea that the working field of view is covered by an image with limited distortion relative to a rectilinear image projection, and assumes that the image dimension on the monitor is proportional to the viewing angle although it is not precisely true.

The term “of interest”, “orientation of interest” or “dimension of interest” are often used to indicate a particular orientation or dimension to be considered or worked on.

2 Modulation Transfer Function, Definition of MTF10 and MTF50

Visual information displayed through a CMS is a result of several sequential information transfers starting from a reflection of the environmental light by the target scene up to the output image on the monitor where a driver visually observes the image according to his visual acuity. The ability to resolve and distinguish details is determined by a number of intermediate factors composing the CMS. It is not the scope of this handbook to go into details of each of these factors, but rather to evaluate the integral results of all these optical, optoelectronic, electronic processing and transfer of the visual information to the driver. To get a brief view, Fig. 1 illustrates several of such steps and consequential effects that may occur in the image transfer flow. The evaluation is performed by capturing black and white luminance modulation of a test chart containing a varying spatial frequency and analysing the transferred visual output on the monitor using a reference camera, by evaluating multiple spatial frequencies of the reproduced test chart on the monitor.

The capability of imaging equipment to capture and transfer an image through the system is generally dependent on the spatial and temporal frequency. The modulation transfer capability is a function of the spatial frequency and it is

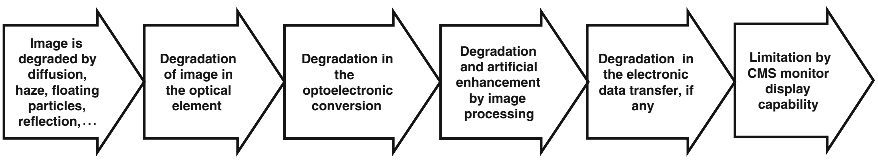


Fig. 1 Image is affected in the transfer flow

generally a decreasing function of spatial frequency. The MTF being a function of frequency, can be denoted as MTF(f) and is given as a modulation signal normalized to modulation signal at 0 frequency; or MTF(0) = 1, where f is argument representing the spatial frequency. For details, see Eq. (13) in Sect. 8.3. In ISO 16505:2015, the evaluation of MTF shall cover not only the capturing element (camera) but also the display element (monitor) performance.

A simple method for evaluating the MTF is by using a resolution charts composed of black and white chart with variable spatial frequency. An example of a suggested chart in ISO 16505:2015 is given in Fig. 5 containing hyperbolic wedged lines to create a periodic pattern with variable frequency. The chart is captured by the CMS camera and displayed on the monitor. The chart shall also include a reference black and white level nearby the hyperbolic lines, which is used to obtain the representative modulation level for the zero spatial frequency. The output image on the CMS monitor is then captured using a reference camera and the luminance intensity modulation is evaluated. Multiple spatial frequency points of the chart are measured to obtain the MTF(f) curve, starting from low frequency and gradually moving to higher frequency part of the chart.

From this MTF(f) curve, a specific spatial frequency point where the MTF (f) value drops to 10 % is taken as the representative spatial frequency determining the "limit resolution" of the system under evaluation, and it is defined as "resolution (MTF)" in ISO 16505:2015. The adoption of 10 % criteria is partly based on the idea of the Rayleigh Criteria corresponding to 9 % modulation accepted as criteria to judge whether two luminance spot becomes distinguishable to each other or not, though the precise physical background is not the same. This particular spatial frequency is denoted as MTF10 in ISO 16505:2015 and used as single parameter to define the resolution (MTF) requirement. It should be reminded that ISO 16505:2015 adopted the 10 % modulation as the limiting resolution point instead of the 5 or 3 % value which is often used in the imaging industry. This higher value has been chosen considering the needs of faster perception in the CMS application, compared to a static observation of a captured image in still picture applications.

Another aspect influencing the perception in a CMS is the sharpness. From the same MTF(f) curve, the spatial frequency where the MTF(f) value drops to 50 % is taken as the representative spatial frequency characterizing the sharpness of the system. This particular spatial frequency is denoted as MTF50 to define the sharpness requirement. There are several other aspects like edge acuteness or luminance contrast that also affect the perceptual sharpness.

3 Image Height and Measuring Units

First of all, it is useful to know the different usage of the term "image height" in the imaging industry and optics.

In the imaging industry, the term "image height", or "picture height", is often used to refer to the shorter dimension of a rectangular image whichever the orientation of an image is. The reason why it is called "image height" is given in the following paragraph. An image layout with shorter dimension oriented to vertical is usually called "panorama" layout mode and an image with shorter dimension oriented to horizontal is called "portrait" layout mode.

In the optics on the other hand, the lens proprieties like aberration vary according to the distance from the optical axis and the term "image height" refers to distance from the central optical axis to periphery. In an ideal cylindrical optical design, the optical characteristics are symmetric regardless of the azimuthal orientation and the characterization curve is dependent on this *image radial height*, or simply called "image height". The term "image radial height" is partly used in Annex G of ISO 16505:2015, to differentiate it from the image height, that is, the image shorter dimension.

The electronic imaging industry started its history from the television system with a long history using image like 4:3 aspect ratio, or 16:9 aspect ratio with the later introduction of high definition television system. Generally a television image is viewed as a rectangle image with its shorter dimension in the vertical orientation in its panorama layout. And it was quite natural to call this vertical dimension as "image height". This image height is very practical to be used as a reference dimension in characterizing the resolution of a raster scan television system because it can be used independent of the display size and its interpretation could be expanded to horizontal resolution by referring to image vertical height as well. The electronic imaging departed from the traditional television system using the panorama layout, allowing the same system to be used in 90° rotated condition and in such case, the reference "image height" is no longer the vertical height but *the shorter dimension* of an image in evaluation instead. Today, the electronic imaging industry expanded the system resolution evaluation to a variety of new application other than television; but it keeps using the image height as reference to define the resolution because it can be used independent of image size, and the reference dimension is self-contained in the image. A deviation of the image aspect ratio from the standard format may directly influence the nominal resolution value of the system when evaluated using [LW/PH] units. We will see these effects with some examples in the next section.

The second edition of ISO 12233 has changed the terms and definitions wording so there are no explicit citations of the "shorter dimension" being the "image height", namely the "PICTURE HEIGHT". But one can trace back to the first edition, ISO 12233:2000 explaining the definition of the units [LW/PH], citing the height to be the active image distance in the shorter test-chart dimension.

4 Specific Use of the [LW/PH] Unit Within ISO16505:2015

The definition of the Line Widths per Picture Height or [LW/PH] unit in ISO 12233:2014 refers to the shorter dimension of an image under evaluation, whichever the orientation of the image is or whatever the image format is. And more precisely, it refers to the full size shorter dimension of the image under evaluation. On the other hand, ISO 16505:2015 defines the resolution (MTF) requirement on *a hypothetical squared image covering a defined monitor size*, or used in partial image size under evaluation. The "defined monitor size" is a terminology adopted in ISO 16505 and it is a region on the monitor that shall be declared by the applicant for the purpose of type approval. Technical Services will use this defined monitor size to calculate and derive necessary parameter for compliance verification of the CMS. This means that the reference "image height" within the scope of the ISO 16505:2015 is no longer the same one used to refer to the full image size. From the needs to differentiate, ISO 16505:2015 adopted an additional subscript text "(1:1)" after MTF to read like $\mathrm{MTF10}_{(1:1)}$. This denotation is a way to explicitly show that the definition is made on a squared assumed image or sub-image. The modification of the evaluation image size keeps the requirement independent of aspect ratio of the full image size displayed on the monitor, and also to the possible magnification aspect ratio if it applies.

We learned that the modulation transfer function is a function of the spatial frequency and when evaluating in units of [LW/PH], the nominal value becomes dependent on the actual evaluation size of the image. To improve the understanding of the dependency of the nominal value on evaluation image size, one may rewrite the $\mathrm{MTF}(f)$ as $\mathrm{MTF}_{(1:1)}(f,\ W_{evaluation\_image\_size})$; or simply denote as $\mathrm{MTF}_{(1:1)}(W_{evaluation\_image\_size})$, thus visualising the evaluation image size as an arguments influencing the nominal spatial frequency value. This visualisation helps to follow the dependency of the nominal spatial frequency value given in [LW/PH] units according to different evaluation image size, because the number of distinguishable lines is linearly proportional to this image size used in the evaluation.

In the evaluation of conventional resolution of electronic still picture image, such a denotation of the dependency on image size is not necessary when following ISO12233:2014 because the spatial frequency measured in [LW/PH] is always measured as full image size and it automatically references the fixed shorter dimension of the full size image, by convention.

Another description particular to ISO 16505:2015 is given in the Sect. D.4 of ISO 16505:2015. When a CMS is measured according to the procedure determined by ISO 12233:2014, conversion will be necessary to convert the requirement into the image format with determined size and aspect ratio. MTF subscripted with text "(W:H)": like $\mathrm{MTF}_{(\mathrm{W:H})}$ is provided to describe the nominal resolution requirement according to specific horizontal and vertical aspect ratio of W:H of particular design. The information is given to keep consistency and not necessarily as a suggested procedure to be followed.

Now, what does Line Widths per Picture Height, namely [LW/PH] unit exactly means? It is defined as number of countable line widths that could fit into the shorter dimension of an image, where a black line is one "width" and the corresponding interval white line is another "width". From this definition, an image with nominal resolution of N [LW/PH] may contain no more than half number of black and white line(s). We deal with spatial frequency and it might look strange not to count the black and white pairs as one cycle. But the idea of counting the lines composing an image has historical reasons. When television system were the main target of the resolution evaluation of electronic imaging system, the number of distinguishable raster scan lines was one of main reference in the resolution measurement of such systems. This definition of Line Widths per Picture Height shall not be confused with the unit defined as Line Pairs per Picture Height or Cycles per Picture Height, which gives half the value.

There are other spatial frequency units that are image size or format invariant like line pairs per millimetre or cycle per millimetre which are often used in the evaluation of optical components. But these units are not suitable for the particular purpose of evaluating the integral performance of a CMS and it was not adopted in ISO 16505:2015.

When considering a requirement for CMS, referencing the full size shorter dimension of an image may create confusion and it does not help in understanding the CMS resolution (MTF) requirement. Therefore, the information provided here for ISO 12233:2014 could be initially skipped until the CMS requirement is well understood.

The nominal spatial frequency value in [LW/PH] units is flexible, if not to say fragile and insignificant, until the precise image size and format is determined. To understand the problem, see the illustrative examples in the Fig. 2 where the physical spatial frequency is same in all images.

The Fig. 2a shows an illustrative example of an image with the size Horizontal × Vertical = 400 (mm) × 300 (mm). The nominal horizontal spatial frequency of black and white line periodic image is 30 [LW/PH] as it contains 30

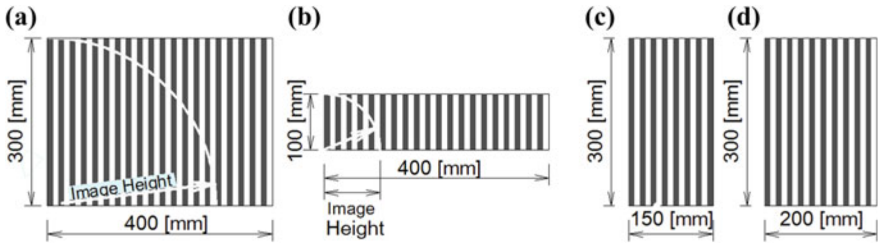


Fig. 2 Illustration of different nominal horizontal resolution values of different images with same physical spatial frequency but evaluated in different image formats. **a** 30 [LW/PH]. **b** 10 [LW/PH], **c** 15 [LW/PH], **d** 20 [LW/PH]

line widths when we fit them into the reference image shorter dimension, namely its "PICTURE HEIGHT" which is illustrated by an arc. But when the same image containing same physical spatial frequency bars evaluated in an image size where the shorter dimension is cut by a third to become 400 (mm) × (300/3) (mm), it will now read its nominal spatial frequency to be 10 [LW/PH], because the reference image "HEIGHT" is 1/3 or the original dimension of the image in Fig. 2a. Figure 2c, d are different conditions in which the "PICTURE HEIGHT" no longer refers to the original vertical dimension but refers to the horizontal orientation (shorter dimension), thus causing the nominal value of spatial frequency to be proportional to the image horizontal size, in a range of up to the horizontal width equal to 300 (mm). Again, all images contain periodic bars with exactly the same physical spatial frequency.

Working with such a variable nominal resolution value may only create a lot of confusion with little benefit to the characterization of a CMS if used according to ISO 12233:2014, with units based on full size image. But on the other hand, resolution measurement described in ISO 12233:2014 is a result of long history, and this standard provides a well-established measurement procedure and units. The preference was to create the CMS requirement in a way to make the requirements independent of the design of the output image size of a CMS.

To cope for the above problem, ISO 16505:2015 has adopted a definition so that all requirements are based on *a hypothetical squared image covering a defined monitor size* image. The adoption of such hypothetical squared image makes the requirement neutral to the final image format, irrespective to final displayed image aspect ratio and further magnification aspect ratio. To understand the requirement of ISO 16505:2015 one needs to take into account that the dimension of interest is not the shorter dimension of the entire image, but instead, it is the dimension in the orientation of interest which is created by the hypothetical squared image covering the defined monitor size. Or more basically, the dimension determined by the required minimum field of view, although this is not explicitly used for a different reason. Whenever applying ISO 12233:2014 standard's method by means of evaluation in the full size image, it has to be converted according to each specific system design.

Again, it is important to keep reminded that the nominal spatial frequency value using units of [LW/PH] is image size and image format dependent. When evaluating in a partial area of the original image, the number of countable line widths within the partial image is proportionally reduced and the nominal resolution value shall be recalculated so that it refers to this new partial image height. This size correction factor according to image size will be necessary because the CMS image will be displayed in part of the monitor, and not necessarily covering the full size of the output image. There is no explanation of such a correction in ISO 12233:2014 because the "Picture Height" in principle refers always to the full size shorter dimension of an image under evaluation, independent of its image orientation. This is because ISO 12233:2014 does not target to measure the characteristic of a partial image, or if necessary to quantify in a size independent way, other unit like [LW/mm] is also available for such a purpose.

5 CMS Resolution (MTF) Requirement

In principle, the evaluation of limit resolution is intended to primarily verify the spatial frequency with resolution values measured in units of [LW/PH]; it corresponds to the number of distinguishable lines which are observable within a defined area. In ISO 16505:2015, the spatial frequency is defined and evaluated in a squared image in a defined area. This defined area being a squared image; it means that when measuring the horizontal resolution, the shorter dimension is equal to horizontal dimension itself. The explanation and equation are given only for the definition of the horizontal resolution of the CMS in this chapter. To obtain the equations for the vertical resolution, simply interchange the width "W" with the height "H" and the suffix "hor" with "ver".

The resolution (MTF) requirement for CMS is defined in a hypothetical squared image, with horizontal and vertical dimension which is determined by the assumed monitor defined size. The nominal resolution value is proportional to the defined dimension "monitor defined size". This is the area on the CMS monitor that shall be determined and declared by the applicant for type approval. The nominal resolution value given in [LW/PH] uses the virtual height of the image. This image being a squared image by definition makes the shorter dimension to be always equal to the horizontal dimension in the orientation of interest because it is a square image. Therefore, the image height does not necessarily show up in the equation.

The basis of this resolution (MTF) requirement can be traced back to verify the number of distinguishable line widths that a driver can observe within the required minimum field of view, whose angle is defined as $\alpha_{mirror/hor/min}$.

Now, the visual acuity V_{eye} of a human eye is defined such that its inverse value $1/V_{eye}$ gives the minimum angular size, in arcmin unit, of what the subject person can observe and distinguish with his naked visual acuity. And when observing a scene through a path in which the scene is magnified by a magnification factor of the convex mirror M_{mirror}, the actual angular size that the subject person can actually observe becomes $1/M_{mirror}$ of the $1/V_{eye}$. Under present regulation, a driver with minimal visual acuity can drive using a traditional type approved mirror, and the minimum observable angular detail in such combination is calculated by the following Eq. (1).

$$\frac{1}{M_{mirror/avg}} \times \frac{1}{V_{eye/min}} \times \frac{1}{60} \times \frac{^\circ}{arcmin} \tag{1}$$

where $M_{mirror/avg}$ denotes the average magnification factor of a type approved convex mirror, $V_{eye/min}$ denotes the minimum visual acuity allowed for a licensed driver, and the last factors is added to convert arcminutes unit to degrees unit. This is the effective minimum visual acuity that a driver shall at least be able to observe using traditionally mirror.

Therefore, the resolution (MTF) requirement can be calculated as the numbers of distinguishable widths observable within the required minimum field of view and it can be calculated by dividing the $\alpha_{\mathrm{mirror/hor/min}}$ by the above effective visual acuity, providing the nominal resolution value calculated by Eq. (2), if evaluated in a monitor size displaying the required minimum field of view.

$$\begin{aligned} &\mathrm{MTF10}_{\mathrm{MIN}(1:1)/\mathrm{hor}}\left(W_{monitor/hor/\mathrm{min}}\right) \\ &= \mathrm{M}_{\mathrm{mirror/avg}} \times \alpha_{\mathrm{mirror/hor/min}} \times \mathrm{V}_{\mathrm{eye/min}} \times 60 \times \frac{\mathrm{arcmin}}{\circ} \end{aligned} \tag{2}$$

The requirement given as Eq. (15) in ISO 16505:2015 is an expansion of the above equation to nominal resolution requirement given for the required minimum field of view to the monitor defined size of $W_{\mathrm{monitor/hor}}$, where the first factor $(W_{\mathrm{monitor/hor}}/W_{\mathrm{monitor/hor/min}})$ is the size compensation factor to expand the evaluation window from $W_{\mathrm{monitor/hor/min}}$ to $W_{\mathrm{monitor/hor}}$. Considering the dependency of the nominal value to the measured image size, the same Eq. (15) is denoted as shown by the following Eq. (3)

$$\begin{aligned} &\mathrm{MTF10}_{\mathrm{MIN}(1:1)/\mathrm{hor}}\left(W_{monitor/hor}\right) \\ &= \left(\frac{W_{monitor/hor}}{W_{\mathrm{monitor/hor/min}}}\right) \times \mathrm{M}_{\mathrm{mirror/avg}} \times \alpha_{\mathrm{mirror/hor/min}} \times \mathrm{V}_{\mathrm{eye/min}} \times 60 \times \frac{\mathrm{arcmin}}{\circ} \end{aligned} \tag{3}$$

in this handbook.

The required field of view in a CMS is expected to be relatively narrow, at least for class I, II and III and the relevant image area that shall be observable by the driver is covered with an image projection close to a rectilinear projection with relatively small distortion. Under the above condition, the following approximation given by Eq. (4) applies,

$$\frac{\mathrm{W}_{\mathrm{monitor/hor}}}{\mathrm{W}_{\mathrm{monitor/hor/min}}} \approx \frac{\alpha_{\mathrm{monitor/hor}}}{\alpha_{\mathrm{monitor/hor/min}}}, \tag{4}$$

where $\alpha_{\mathrm{monitor/hor}}$ is the field of view which is displayed in the monitor defined size. Then, Eq. (3) can be simplified by the following Eq. (5)

$$\begin{aligned} &\mathrm{MTF10}_{\mathrm{MIN}(1:1)/\mathrm{hor}}\left(W_{monitor/hor}\right) \\ &\approx \mathrm{M}_{\mathrm{mirror/avg}} \times \alpha_{\mathrm{monitor/hor}} \times \mathrm{V}_{\mathrm{eye/min}} \times 60 \times \frac{\mathrm{arcmin}}{\circ} \end{aligned} \tag{5}$$

and the resolution (MTF) requirement can be verified if $W_{\mathrm{monitor/hor}}$ and $\alpha_{\mathrm{monitor/hor}}$ is a known value.

The adoption of Eq. (3) or (5) compared to Eq. (2) creates a disadvantage in terms that the nominal resolution requirement value becomes again variant to evaluation image size and it is not a fixed value. But the adoption of Eq. (5) as requirement avoids the need to clarify the exact area determining the required minimum field of view, which is not available until the vehicle installation design is finalized. Equations (3) and (5) brings more flexibility during type approval test procedure because there will be no need to precisely define the vehicle installation parameters. This will be important to enable verification of CMS for type approval before all design parameters of a vehicle is determined, at development stage for example. Nevertheless, as shown in the examples in Fig. 2, the physical spatial frequency is a constant value and the compensation factor to the nominal resolution value is just an adjustment of the nominal value to the evaluation window size when defining in [LW/PH] units.

When evaluating a CMS of class VI, a large deviation to the assumption given in Eq. (4) may occur especially if the monitor defined size is expanded to an extreme wide angle. CMS exhibiting large peripheral distortion is likely be affected. It is therefore advised to determine the requirement value in the smallest possible monitor defined size, preferably near or equal to the required minimum field of view.

6 Measuring the Spatial Frequency

The conversion guide given in the Annex D.4 of ISO 16505:2015 for different case of aspect ratio is intended to fill the gap coming from the "image height" definition in ISO 12233:2014 and squared image used in ISO 16505:2015. The conversion will be required only if the measurement is performed on the monitor defined size in which the image aspect ratio does not follow a 1:1 aspect ratio. In short, the conversions instructions were provided in ISO 16505:2015 to give consistency to the existing ISO 12233:2014 without re-creating a new measuring unit specific for the CMS. Further complication may occur when involving magnification aspect ratio other than (1:1), involving modification of the reference height when referring to the orthogonal orientation.

One can avoid such conversions by verifying requirements following procedure given in Annex D.5 of ISO 16505:2015, which uses a partially captured squared image in the evaluation. ISO 16505:2015 adopts the Line Widths per Picture Height but restricting the size to a specific squared sub-image. The modification of the evaluation image size means that the actual number of distinguishable lines will proportionally decrease according to the reduced image size. Therefore, Eq. (3) can be further modified to a square cropped image within the displayed monitor image. By correcting the size modification to the cropped image in an evaluation window size of W_{crop}, the nominal resolution value becomes

$$\begin{aligned} &\mathrm{MTF10}_{\mathrm{MIN}(1:1)/\mathrm{hor}}\left(W_{crop}\right) \\ &\quad = \left(\frac{\mathrm{W}_{crop}}{\mathrm{W}_{\mathrm{monitor/hor/min}}}\right) \times \mathrm{M}_{\mathrm{mirror/avg}} \times \alpha_{\mathrm{monitor/hor/min}} \times \mathrm{V}_{\mathrm{eye/min}} \times 60 \times \frac{\mathrm{arcmin}}{\circ} \end{aligned} \tag{6}$$

or to derive from the monitor defined size the approximation applies

$$\begin{aligned} &\mathrm{MTF10}_{\mathrm{MIN(1:1)/hor}}\left(W_{crop}\right) \\ &\approx \left(\frac{\mathrm{W}_{crop}}{\mathrm{W}_{\mathrm{monitor/hor}}}\right) \times \mathrm{M}_{\mathrm{mirror/avg}} \times \alpha_{\mathrm{monitor/hor}} \times \mathrm{V}_{\mathrm{eye/min}} \times 60 \times \frac{\mathrm{arcmin}}{^{\circ}} \end{aligned} \tag{7}$$

The right side of Eq. (6) is equivalent to formula D.6 given in Annex D.5 of ISO 16505:2015.

The Fig. 3 shows an example to visualize the effects of the modification of evaluation window. The image of a black and white parallel bar chart image exhibiting a horizontal spatial frequency equals to 28 [LW/PH] measured according to ISO 12233:2014, which by definition, means that the same number of line widths fill exactly into the shorter dimension of the full image size. But those readers who are not familiar with ISO 12233:2014 might question that we could observe 38 line widths in the horizontal orientation. Now, remember that ISO 12233:2014 defines the spatial frequency relative to the shorter dimension, which in this case is the image vertical width. If we count the number of vertical lines in a horizontal orientation, one finds that 28 lines can fit exactly into the height of this picture image. An arc with the radius equal to the image height can help to visualize it (see the arc in Fig. 2 as an example).

Now, moving to the definition adopted in ISO 16505:2015, the actual picture height of a rectangular image is not directly used as reference anymore. The Eq. (2) is the basis of the CMS resolution (MTF) requirement and the nominal spatial frequency value given in units of [LW/PH] is proportionally dependent on the evaluation image size in its own orientation.

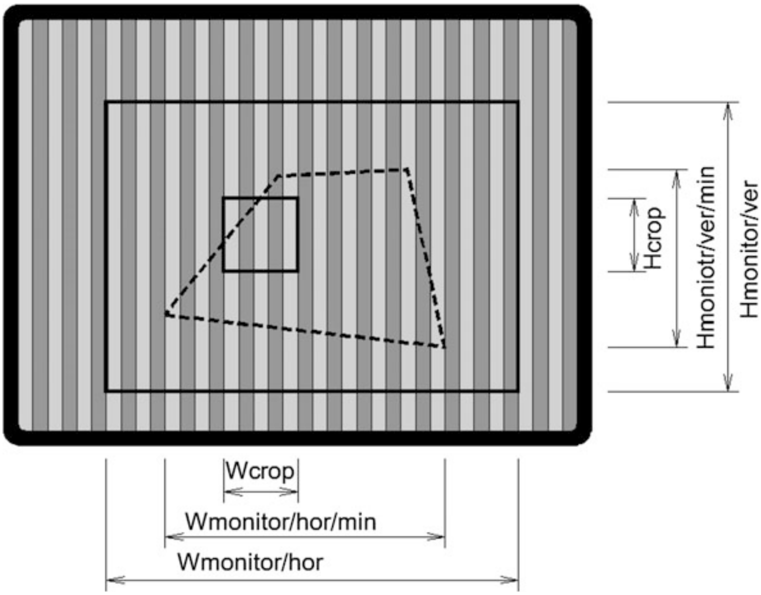


Fig. 3 Correlation of different definition areas

It might look strange to have Eq. (3) or (5) as primary equation defining the requirement in ISO 16505:2015 and keep Eq. (2) secondary but as it was explained in the previous section, this was a compromise to enable the type approval evaluation of CMS devices without precisely defining actual area determining the required minimum field of view. It gives flexibility when the actual region defining the required minimum field of view is not precisely determined, for example at component development stage prior to determining all vehicle installation parameters.

As indicated in the beginning of this section, ISO 16505:2015 provides an option to evaluate the CMS by partial evaluation of the output image. The advantage of evaluation procedure based on the partially cropped image is that the characterization of the image captured by the reference camera can be done solely on this image captured by a reference camera in [LW/PH] units of its own image in the reference camera, independent of the image on the monitor.

To visualize the requirement, let us assume that the black and white parallel bar in Fig. 3 is an image of a chart filled with parallel bars of a spatial frequency representing resolution (MTF) minimum requirement with a modulation of exactly 10 %. The horizontal resolution if measured in the applicant declared monitor defined size of $W_{monitor/hor}$ will be $MTF10_{MIN(1:1)/hor}(W_{monitor/hor}) = 28$ [LW/PH]. This nominal resolution value is just an expansion of the $MTF10_{MIN(1:1)/hor}(W_{monitor/hor/min}) = 19$ [LW/PH] measured in a hypothetical square image enveloping exactly the area covered by the required minimum field of view. Note that both nominal resolutions are describing exactly the same physical spatial frequency. And if the spatial frequency is measured in a partial area determined by W_{crop} in which the size is 5/19 of the $W_{monitor/hor/min}$, the number of distinguishable line widths in this window is then proportionally reduced to 5/19 to give

$$\begin{aligned}\mathrm{MTF10}_{\mathrm{MIN}(1:1)/\mathrm{hor}}\left(\mathrm{W}_{\mathrm{crop}}\right) &= (5/19) \times \mathrm{MTF10}_{\mathrm{MIN}(1:1)/\mathrm{hor}}\left(\mathrm{W}_{\mathrm{monitor/hor/min}}\right) \\ &= 5[\mathrm{LW/PH}].\end{aligned}$$

This value is exactly the number of distinguishable lines that we can observe when taking a partial image sized $W_{crop} \times H_{crop}$.

As the relation given in the above example shows, we can reversely obtain the nominal resolution value from the partially evaluated image, independent on the image size how it is captured by a reference camera, and without any need to care about the aspect ratio as long as it is evaluated in a squared image and compensated for the evaluation window size W_{crop} on the CMS monitor itself. From the independently measured nominal resolution value $MTF10_{MIN(1:1)/hor}(W_{crop})$ on the cropped and captured image by reference camera, the nominal resolution value for required minimum field of view is reversely calculated by the following Eq. (8)

$$\begin{aligned}&\mathrm{MTF10}_{\mathrm{MIN}(1:1)/\mathrm{hor}}\left(W_{monitor/hor/\mathrm{min}}\right) \\ &\quad = \left(\frac{\mathrm{W}_{\mathrm{monitor/hor/min}}}{\mathrm{W}_{\mathrm{crop}}}\right) \times \mathrm{MTF10}_{\mathrm{MIN}(1:1)/\mathrm{hor}}\left(\mathrm{W}_{\mathrm{crop}}\right)\end{aligned} \tag{8}$$

or for monitor defined size, it is reversely calculated by the following Eq. (9)

$$\begin{aligned} & \mathrm{MTF10}_{\mathrm{MIN(1:1)/hor}}\left(W_{monitor/hor}\right) \\ & \quad = \left(\frac{\mathrm{W}_{\mathrm{monitor/hor}}}{\mathrm{W}_{\mathrm{crop}}}\right) \times \mathrm{MTF10}_{\mathrm{MIN(1:1)/hor}}\left(\mathrm{W}_{\mathrm{crop}}\right) \end{aligned} \tag{9}$$

and both equations can be derived from the equation and formula given in Sect. 3.16 and Annex D.5 of ISO 16505:2015.

As previously mentioned, optical resolution characteristics of an optical system generally decrease from the optical centre towards its periphery. The requirements of the resolution (MTF) on the corner points are intended to secure a good visibility and enable the driver to observe the relevant range of the scene with appropriate resolution. Requirements are defined at these 4 additional diagonal corner positions of the output image as representative points to verify the requirements, with target criteria half relaxed to the image centre. The measuring positions are defined at 70 % of image radial height in the diagonal orientation. The horizontal and vertical resolution (MTF) shall be verified at these 4 additional points. In an image with aspect ratio of 4:3, the circle passing these points will be equivalent to verifying the 90 % (88 % to be precise) lateral point of the image and it will be well covering the relevant range of observation.

Figure 4 shows these target as cross marked points for measuring the corner resolution (MTF) of the CMS. Due to the fact that the chart cannot be precisely adjusted to bring the very limiting spatial frequency point of the chart exactly to this 70 % image height position, some positioning error is acceptable. Generally, resolution would gradually decrease according to the distance from the image centre and therefore verification to compliance shall preferably be performed beyond this image radial height, helping to secure that the CMS satisfy the requirement within this image radial height (shown as broken line circle). It is advised to keep records of the exact positon that the limit resolution point was confirmed because the actual

Fig. 4 70 % radial image height positions to evaluate the corner resolution (MTF) in the diagonal points

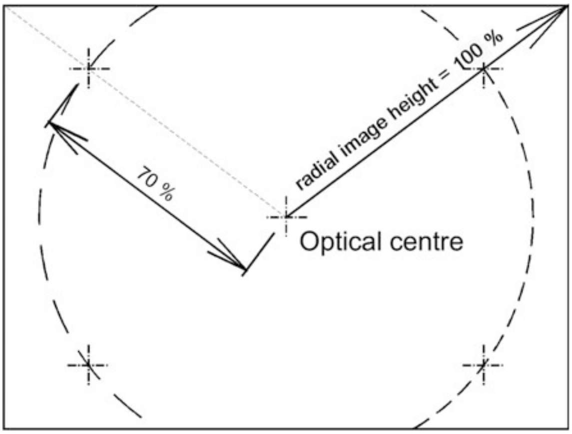


Table 1 Comparison table

	ISO 16505:2015	ISO 12233:2014
Evaluation image format (aspect ratio, dimension)	Hypothetical squared image enveloping the target image. Image aspect ratio is 1:1, in a partial image. It is orientation independent	Measurement on industry standard format, in a full size image format. (e.g. 4:3, 3:2, 16:9)
Resolution (MTF) in [LW/PH] on an square image area enveloping exactly the required minimum field of view, $W_{monitor/hor/min}$ $MTF10_{MIN(1:1)/hor}\left(W_{monitor/hor/min}\right) = M_{mirror/avg} \times \alpha_{mirror/hor/min} \times V_{eye/min} \times 60 \times \frac{arcmin}{\circ}$		Adapting the CMS requirement needs correction according to the display image format (namely how the CMS image is displayed) according to guidelines provided in Annex D.4 of ISO 16505:2015. Therefore, a nominal spatial frequency value in [LW/PH] units cannot be uniquely defined as requirement
Resolution (MTF) in [LW/PH] on an square image with monitor size defined as $W_{monitor/hor}$ $MTF10_{MIN(1:1)/hor}\left(W_{monitor/hor}\right) = \left(\frac{W_{monitor/hor}}{W_{monitor/hor/min}}\right) \times M_{mirror/avg} \times \alpha_{mirror/hor/min} \times V_{eye/min} \times 60 \times \frac{arcmin}{\circ}$ $\approx M_{mirror/avg} \times \alpha_{monitor/hor} \times V_{eye/min} \times 60 \times \frac{arcmin}{\circ}$		
Reference "PICTURE HEIGTH" of the [LW/PH]	The reference width is always the evaluation image width in the orientation of interest	"PICTURE HEIGHT" = Image shorter dimension
Unit used to describe the spatial frequency	[LW/PH] is relative to specific square image. It is a unit adapted according to different image under evaluation	[LW/PH] is relative to full image "HEIGTH"
Nominal resolution (MTF) value in [LW/PH] for a determined spatial frequency	The nominal value given in units of [LW/PH] value is proportional to evaluation image size but neutral to image format (magnification ratio) and magnification aspect ratio	Dependent of image output format. Affected by magnification aspect ratio
Limiting resolution criteria	Spatial frequency point where MTF(*f*) equals 10 % or the spatial frequency where the chart periodicity is lost	Spatial frequency point where MTF(*f*) equals 5 %
Evaluation stage in the capturing, process, transfer and display chain	Evaluation of image as displayed on the CMS monitor using a reference camera. (with low pass filtering, to make driver visual acuity equivalent)	Evaluation using the intermediate numerical signal
Measuring chart distance and focusing	Actual use distance with fixed focus	Adjusted to fit the chart and individually focused
Usage of term "horizontal resolution" and " vertical resolution"	According to physical observing orientation	"Hor. resolution" = resolution in the longer dimension and "ver. resolution" = resolution in shorter dimension

verification point depends on the chart positioning and the results of verification may become ambiguous for later analyses if no record is kept.

Table 1 provides a comparison between ISO 16505:2015 and ISO 12233:2014. There are several aspects to be considered when evaluating the spatial frequency of the CMS. Besides it, it is worthwhile mentioning additional care when working with CMS.

As already adopted in the equations given in this chapter, the denotations of nominal resolution according to different evaluation window size are newly added in this handbook.

7 Evaluation Procedure Depending on the System's Performance

The CMS is evaluated as an integral system including the display performance. It shall consider evaluating how the output image is visually perceived by a driver. A reference camera is used to capture the output image and evaluate the resolution of the CMS. To avoid complexity of the evaluation, ISO 16505:2015 measures the average luminance signal and does not go into details of subpixel design, difference of sensitivity according to colour spectrum or further effects of motion resolution but rather evaluate the CMS in a simplified condition in neutral colour, using black and white hyperbolic chart with variable spatial frequency.

Capturing of image by CMS camera is performed by sampling the image projected on the sensor plane by an array of photosensitive pixels and then displayed on the monitor by a two dimensional array of pixels. For the resolution (MTF) and sharpness evaluation of the CMS systems, a reference camera is used preferably with sampling frequency at least 4 times higher to the pixel composing the CMS. It is advised to avoid the use of sampling frequency lower than twice of the CMS array frequency because it may result in occurrences of considerable aliasing and creating possible spurious signal influencing the results of the measurement. Note that the reference camera does not necessarily need to capture the entire monitor defined size, and the partially captured image shall follow the above recommendation.

Technically one can design a large monitor with large unit pixel design in which the layout details of each pixel design become apparently well observable. But such a large scale design will require a large monitor size to show enough field of view of the CMS, and consequently obstructing the direct vision of the driver to cover for the expected field of view. Although ISO 16505:2015 does not limit the monitor maximum size, it is expected that a CMS will be designed in such manner that the monitor composing the system applies a pixel design small enough; with the details of the pixel design not visually observable unless magnified. On the other hand, even the recent state of the art display technologies in which the individual pixel design is no longer observable by a naked eye, the actual microscopic design is composed by subpixel of effective light emitting area of primary colours surrounded by black screen window. Due to such pixel design, artefact is often observed on projected images on large screen using a video projector, exhibiting what is called screen-door effects as an example. Or one may also have experienced visually observing an expanded pixel when a water droplet fall onto a monitor screen or observing a monitor with a magnification loupe, making the frames and primary colour of subpixel well observable. These microscopic details of the pixel design might be observed during the evaluation of the output image using a reference camera, because it is suggested to use a higher spatial sampling. If not properly filtered to average these high frequency details, the output signal of the reference camera may exhibit high signal peak on active light emitting area and valley on the subpixel dark frame. But this high peak modulation created by the high frequency sampling is not what a driver observes and therefore the signal must be properly filtered so that the averaged signal equivalent to what a driver observes is actually measured.

The resolution is checked by verifying the modulation of the output signal starting from low frequency signal and moving to high frequency signal with the periodic output signal of the chart correctly reproduced on the output image. The point where the modulation drops down to 10 % or the point where the correct periodicity gets lost is taken as the limit resolution of the CMS. Details of measurement procedure will be given later. There may exist CMS in which even after averaging of the high frequency signal, some large abnormal signal remains on the output image, and the correct periodic signal of the 5 black lines will not be correctly reproduced when checking from low spatial frequency to higher frequency pattern of the hyperbolic chart before the MTF(f) drops to 10 %. The irregularity of the periodic signal occurs due to the aliasing and further combined with effects of the signal processing to enhance high frequency edge of an image to artificially create a sharper image. As a result, correct periodicity of the chart might be lost before the signal modulation drops to 10 %. For such system, the limit resolution MTF10 is defined as the spatial frequency in which this irregularity becomes apparent, replacing the 10 % criteria.

Due to cost and preference to use larger pixel design process from reliability perspective, it is expected that at early stage of the CMS market deployment, the CMS will be using components with pixel counts (number of addressable photo-elements) near the minimum needs for achieving minimum resolution requirement. In the evaluation of such CMS, the signal obtained by oversampling using a reference camera may exhibit large amount of aliasing and residual side effects of the CMS itself, and the capability of reproducing the spatial frequency will be limited by these design constraint rather the optical performance of the camera composing the CMS. In these systems, the anomaly of the periodicity in the output image becomes dominant making it difficult to determine a spatial frequency point where the output image signal drops to a modulation level of 10 %. To cover such system, ISO 16505:2015 adopts an auxiliary criterion for judging whether the CMS is still acceptable as compliant to minimum requirement in terms of distinguishability of details of the scene. The evaluation of the CMS in such case uses an auxiliary parallel bar chart with spatial frequency equal to the required minimum resolution (MTF) $MTF10_{MIN(1:1)/hor}$. The output image is then compared to image given as guideline of acceptance under ISO 16505:2015 Annex D.3.

In brief, the compliance verification of the CMS shall take different procedure according to the performance of the CMS either by

1. Using the hyperbolic chart verifying the spatial frequency point where the modulation transfer function drops to 10 % or
2. By visual inspection using the auxiliary parallel bars chart with spatial frequency equals to $MTF10_{MIN(1:1)/hor}$.

8 Measurement of Resolution and Sharpness

8.1 Single Hyperbolic Chart Versus ISO 12233:2014 Standard Chart

Horizontal resolution is typically measured using a vertical or near vertical oriented test chart feature, and vertical resolution is typically measured using a horizontal or near horizontal oriented test chart feature. An example of a chart to check the horizontal resolution is shown in Fig. 5. The same chart shall be rotated by 90 for the verification of the vertical resolution.

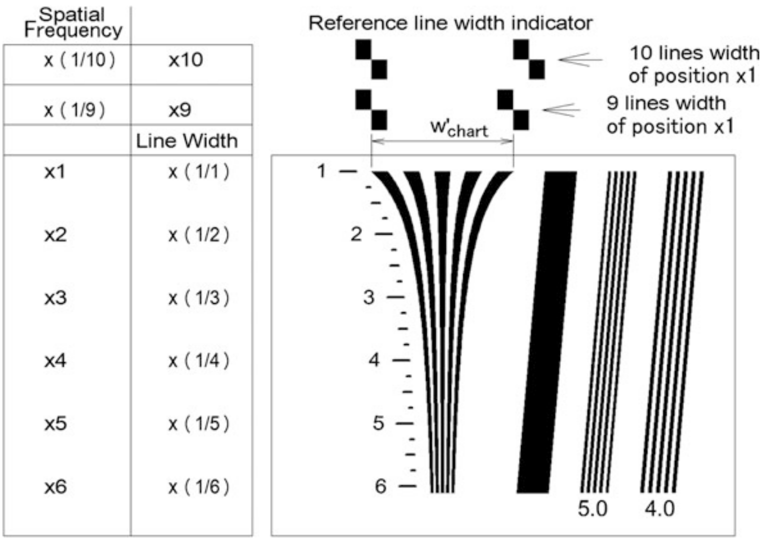


Fig. 5 Hyperbolic chart for resolution measurement and auxiliary parallel bar chart

The camera in a CMS is expected to be a fixed focus type with best focusing range around several meters away from the CMS camera position. Therefore, unlike a conventional still photography camera or imaging system, the CMS resolution cannot be evaluated at close distance by using small standard chart and re-adjusting the lens focus to the evaluation chart position at proximity distance.

To evaluate the system in an actual work distance several metre away from the camera, it will require a huge chart if applying a standard chart design like those defined in ISO12233:2014. And every system has different field of view; and it requires preparing a custom made huge resolution charts for every new design. Creating such a huge chart is technically possible but it is not practical in real use. A practical way to evaluate the system is by splitting the evaluation in several regions using a partial image rather than using a single chart covering the entire image of the CMS. When using a partial hyperbolic chart, the inconvenience will be the missing reference chart height to evaluate and quantify the resolution in units of [LW/PH]. In ISO 16505:2015, an alternative reference width is prepared nearby the measuring position of the hyperbolic chart to substitute for the chart height.

This section provides some instructions on the preparation of the test chart using a single hyperbolic chart and a guide to work with it. The step-by-step instructions given in ISO 16505:2015 differs from a conventional adjustment procedure of resolution measurement given in ISO 12233:2014, in which the chart distance to the camera is adjusted so that the captured image by the camera has its longer dimension outermost line of the image fitting exactly to the chart height. Step (3) described in the test procedure in Sect. 7.5.3.1 of ISO 16505:2015 can be interpreted as a calibration step to determine the spatial frequency at position 1 of the chart, instead of automatically determining the spatial frequency by adjusting the standard chart to the full image height in ISO 12233:2014. Consequently, all numbers which are printed adjacent to the hyperbolic chart will have different significance and its meaning will be explained in the next section.

8.2 Introduction of the Single Hyperbolic Chart

Several different resolution charts are available on the market according to different industry standards. Hyperbolic resolution test chart is one of the test pattern defined under ISO 12233:2014. These charts take their own full size height to define the resolution in [LW/PH] units. In ISO 16505:2015, a single hyperbolic chart similar to the one included in the ISO 12233:2014 is used but the design is slightly different in the point that the widest line width of the hyperbolic lines itself is taken as reference of width. In a conventional chart, the index numbers provided adjacent to the hyperbolic chart are numbers intended to automatically and directly show the spatial frequency of the corresponding point on the chart; or given in number as multiple of 100 when the chart height is adjusted to the image height. Some hyperbolic resolution chart might start with the lowest spatial frequency other than 1 or 100, according to different needs to verify different range of spatial frequency.

A standard resolution chart takes the shorter dimension of the chart as a reference of unit given as line widths per picture height, or [LW/PH]. The evaluation is a self-consistent evaluation where the image resolution is calculated relative to image own height. The value given adjacent to the hyperbolic chart shows the spatial frequency of that specific position, relative to the reference image height. For example, a chart position showing 200 has a line width equivalent to 1/200th of the full image height and its spatial frequency is 200 [LW/PH], which means that 200 lines of a width at that position can fit into the picture height. The benefit of such unit definition is that the evaluation can be performed in any type of standard image regardless of the actual image size, or observation distance.

On the other hand, the numbers shown adjacent to the proposed test chart of ISO 16505:2015 are number intended to show the multiplication factor of spatial frequency relative to the spatial frequency at reference position 1. In the chart given in ISO 16505:2015, the line width of the hyperbolic chart uses the widest width as reference (with a multiplication factor equals to one at this position 1) and the spatial frequency at different positions can be calculated by multiplying the multiplication factor to the spatial frequency at position 1.

For those who are not familiar how a hyperbolic chart design is calculated, following explanation may help to understand it. Mathematically, the width of the line composing the chart is determined as a reciprocal function (or a multiplicative inverse function) to its position. Therefore, the width of the line in these charts decreases reciprocally according to its position and the spatial frequency increases proportionally to its position. If we denote the chart line width at chart position 1 to be $W_{line}(FMF = 1)$ where FMF stands for the frequency multiplication factor, the chart line width at different frequency multiplication factor FMF will be calculated to be $W_{line}(FMF) = W_{line}(FMF = 1) \times (1/FMF)$. As an example, the point with FMF = 4 will contain periodic lines with 4 times higher spatial frequency compared to the spatial frequency at position 1 and the width of the lines will be 1/4th compared to the width at reference position 1.

In ISO 12233:2014, the primary reference dimension in the evaluation using [LW/PH] units is the “PICTURE HEIGHT” of that specific image. However, in the resolution characterization of a CMS according to ISO 16505, the target of the evaluation is to verify the amount of distinguishable information observable in its own orientation of interest, and not necessarily referencing to the image height as is the case in ISO 12233:2014. The dependency of the nominal resolution value according to image format makes it difficult to define a universal requirement value for the CMS and it causes difficulties in correlating the requirement value to the traditional way of denoting the spatial frequency in units of line widths per picture height. In view of this problem, ISO 16505:2015 has adopted the idea of creating a hypothetical square image to define the resolution and sharpness related requirements. This square image makes the issue of image height neutral to orientation, image aspect ratio or magnification aspect ratio. All evaluation related to resolution

(MTF), sharpness and depth of field shall be defined and handled in this hypothetical squared image.

In the absence of chart height, the chart now needs a substitutional dimension to determine the spatial frequency. One can try to measure a single line width at the positon 1 directly on the chart but once the chart image is transferred through the system, the accuracy to read the width of a single line becomes drastically difficult and making the evaluation unreliable unless using a reference width indicator with easier reading characteristic. Therefore, a reference width indicator to give better accuracy than reading the width of a single line is needed. And a bundle of several line widths is suggested to provide improved reading accuracy. When creating an original chart, a symmetric checkerboard pattern to indicate this 9 or 10 line widths like shown in Fig. 5 is advised because it is robust to image blurring or erosion by adjusting or measuring from centre to centre point of the cross pattern and obtain the w'_{chart} value.

8.3 Spatial Frequency of the Chart at Reference Position 1 and Measurement Procedure

In the resolution evaluation of CMS, the primary dimension used as reference is the dimension in the orientation of interest, and not necessarily the image height of the full image. Defined as a square image, the picture height becomes neutral to image orientation and the dimension in the orientation of interest can be directly referred in defining the resolution (MTF) and sharpness in [LW/PH] units, by taking its own dimension as the virtual image height. All resolution (MTF) or sharpness related parameters given in units of [LW/PH] within ISO 16505:2015 subscripted with the text "(1:1)" are defined in a hypothetical square image, and they shall be verified according to that specific square image size used in defining the requirements for the respective parameter value.

For the measurement of the horizontal resolution or sharpness, a vertically stretched hyperbolic chart like shown in Fig. 5 is captured by the CMS camera and displayed on the CMS monitor. This chart does not have a reference chart height but instead an indicator of width to accurately calculate the width of a single line at position 1, by indicating the width of the total of 9 visible lines. The spatial frequency of the chart position 1 given in units of [LW/PH] is defined as the number of lines of that specific width that fit into one side dimension equivalent to the image height. As previously explained, it is hard to accurately obtain the width of a single line at position 1 and therefore a bundle of 9 lines is used to derive the width of one line at positon 1. This procedure correspond to step 3) under Sect. 7.5.3.1 of ISO 16505:2015, providing the results of Eq. (37) of the same standard.

The square image size used to define the requirement is the defined monitor size $W_{monitor/hor}$ and the width of one line at position 1 is ($w'_{chart}/9$). Thus, the nominal spatial frequency at position 1 in this defined monitor size is calculated by the following Eq. (10).

$$K_{hor}\left(W_{monitor/hor}\right) = \left\{W_{monitor/hor}/\left(w'_{chart/hor}/9\right)\right\}. \quad (10)$$

We learned in Sect. 8.2 that the index numbers given adjacent to the hyperbolic chart of Fig. 5 are abbreviations of the spatial frequency multiplication factors relative to the spatial frequency at position 1 of the same chart. Now that the spatial frequency at position 1 becomes known, all other spatial frequency on the chart can be derived in respect to this reference by calculating the multiplication factor. To obtain the point on this chart determining the required minimum resolution, divide the nominal spatial frequency $MTF10_{MIN(1:1)/hor}(W_{monitor/hor})$ by the nominal spatial frequency at position 1 as given in the following Eq. (11).

$$P_{hor}\left(W_{monitor/hor}\right) = \left\{MTF10_{MIN(1:1)/hor}\left(W_{monitor/hor}\right)/K_{hor}\left(W_{monitor/hor}\right)\right\} \quad (11)$$

Following are procedures to obtain the modulation transfer function. The output image on the monitor is captured by a reference camera as luminance signal and measured using an oscilloscope or image digitizer to obtain a plot of the luminance intensity in the measuring orientation. The range of the plot shall cover the reference white and black region adjacent to the hyperbolic chart representing the zero spatial frequency which will be needed in the calculation of the MTF(f). Modulation transfer function curve can be obtained by measuring multiple point with different spatial frequency starting from lower frequency and moving to higher frequency. The observation from lower frequency to higher frequency is desirable because an evaluation of a single higher frequency point may not provide a correct result. It is because some abnormality in the periodic signal may occur in an intermediate frequency as effects of aliasing and image enhancement. Apparent periodic signal at higher frequency could be just a result of aliasing. A precise MTF(f) curve is not required as verification procedure of CMS resolution measurement and measuring some specific point shall be enough to check the compliance with the requirement as long as it is not a spurious result of aliasing.

The luminance signal of target spatial frequency is measured to obtain the modulation M(f) of the black and white line, and then compared to reference zero frequency modulation $M(f = 0)$ obtained from the white and black area prepared as reference adjacent to the hyperbolic line chart. The small upper figure in the right side of Fig. 6 shows a cropped image from the left image with a further small horizontal stripe at the position where the multiplication factor shows a number about ×3.2. The lower graph is a plot along the above stripe and it shows the luminance intensity periodic signal of this window from the hyperbolic chart, with further white level and black level at its right side which is used to obtain the reference zero frequency modulation.

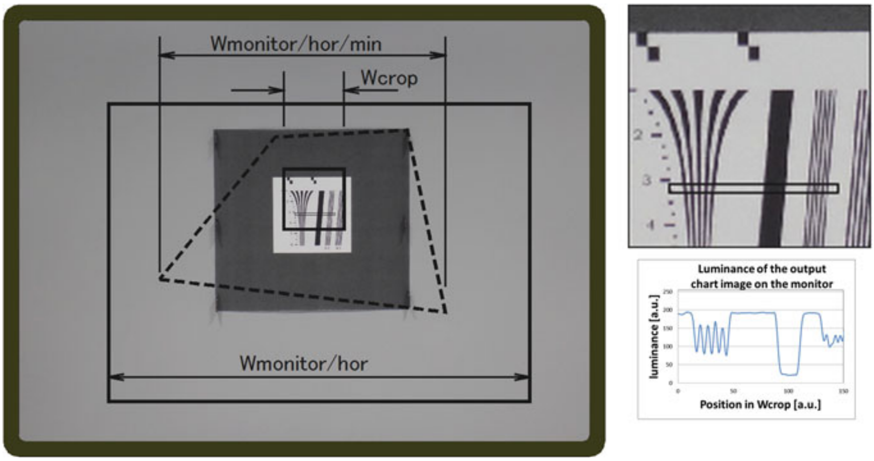


Fig. 6 Illustrative image of a chart reproduced on the monitor and an example of luminance signal of specific horizontal line within a cropped image

The modulation M(*f*) of output image is defined by Eq. (12).

$$M(f) = \{(I_{max}(f) - I_{min}(f))/(I_{max}(f) + I_{min}(f))\}, \tag{12}$$

where $I_{max}(f)$ and $I_{min}(f)$ denotes the luminance intensity at white peak level and black valley level of the specific spatial frequency *f*. The modulation transfer function is a normalized value relative to the modulation at zero frequency and it is calculated by Eq. (13).

$$MTF(f) = M(f)/M(f = 0), \tag{13}$$

where M(f = 0) is the modulation at zero frequency obtained from the spaced white and large black region on the chart. The modulation transfer function is generally a decreasing function to the spatial frequency. Therefore, there could be two approaches in the evaluation of the system according to the needs of the evaluation.

1. For the verification to the compliance according to minimum requirement according to ISO 16505:2015:
   Verify the modulation transfer function at the point defined as the required minimum resolution (MTF). If the modulation transfer function at the spatial frequency $MTF10_{MIN(1:1)/hor}(W_{monitor/hor})$ is over 10 % at this frequency, the actual spatial frequency where modulation drops to 10 % is expected to be higher compared to the above spatial frequency thus satisfying the requirement. Depending on CMS, there might be cases in which limit resolution is limited by the number of pixel count, or image processing causing anomaly to the periodic signal before the modulation drops to 10 %. In such case, an auxiliary parallel bar chart containing line and space patterns of spatial frequency equals to the

minimum requirement $MTF10_{MIN(1:1)/hor}(W_{monitor/hor})$ shall be used to verify whether the output image satisfy the minimum visual criteria given in Annex D.3 of ISO 16505:2015.

2. For the evaluation to find the actual limit resolution (MTF):
The actual spatial frequency where modulation drops to 10 % must be directly verified. In this case, the output response of the image is evaluated from lower frequency to higher frequency until a point where the modulation transfer function drops to 10 %. The spatial frequency of that specific point where the modulation drops to 10 % must be reversely calculated, by multiplying the multiplication index of that specific point on the chart to the spatial frequency of the chart at position 1. If the limit resolution is determined by the optical performance, it is likely that the modulation points where the modulation transfer function drops to 10 % will be directly found by this procedure. However, if the optical MTF performance is maintained high up to high frequency, the limit resolution might be first affected by the sampling limit and/or further image processing causing an anomaly of periodic signal as cited before. The target of the evaluation in this case might not be so critical as to verify the limit to compliance, therefore a simple visual verification by finding the point where an anomaly of the periodicity start occurring can be used to estimate rough limit resolution. If a precise verification of the very limit is needed, the use of a parallel bar chart with spatial frequency close to this rough spatial frequency is advised. Further minor adjustment of spatial frequency could be achieved by changing the distance of the prepared chart to find an exact limiting point when compared to the visual minimum acceptance criteria given in Annex D.3 of ISO 16505:2015. The actual spatial frequency shall be corrected according to the modified final positioning of the chart.

The chart shall be oriented in such a way to bring the resolution hyperbolic chart plane parallel to the orthogonal plane of the optical axis of the CMS. There could be special case in which the displayed image is not a rectilinear projection image according to the optical axis of the CMS camera. For example, if the CMS is composed by a camera with fish-eye type wide field of view lens and CMS displaying a distortion converted image with target orientation out of the camera original optical axis, the chart shall be aligned according to the intended viewing target orientation, and not to the original optical axis of the CMS camera.

Figure 6 shows an illustrative example of full display containing an output image of CMS. The image of a CMS could be a part or the entire display area according to the design concept of the manufacturer but a specific area shall be declared by the applicant for type approval purpose, which is defined as “defined monitor size” ($W_{monitor/hor}$) in ISO 16050:2015. All verifications to compliance are evaluated using this defined image area. Image that could be shown in between the monitor defined size and the entire display area are arbitrary but advised to be a content of low influence to avoid distraction to the driver.

The resolution (MTF) or sharpness is measured in the centre and in the four corner region of the image within the area defined as monitor defined size. Each region of the image is captured by positioning the single chart according to the evaluation target position. The background or optionally the foreground surrounding area shall be covered by a neutral grey plane chart to keep the correct exposure operation of the CMS with a uniform illumination. The following procedure gives instruction to evaluate in a partial image which is provided as informative information in Annex D.5 of ISO 16505:2015. The resolution is evaluated in a partially cropped square image captured by a reference camera. As the nominal spatial frequency value in [LW/PH] units is proportional to the evaluation image size, the nominal resolution value shall be appropriately reduced by the same ratio $W_{crop}/W_{monitor/hor}$, thus deriving the requirements as given in Eq. (6) or (7).

The small square image in the Fig. 6 shows an example of a partial area to be captured by a reference camera for the detailed evaluation using a reference camera. The positioning of chart shall be made such that the appearance of the w'_{chart} is clearly duplicating the reference width nearby the hyperbolic chart in the reference camera image. First, the spatial frequency of the reference position 1 is calculated with the following Eq. (14), in which the captured image window dimension W_{crop} is divided by the width of a single line ($w'_{chart}/9$). In other words, it indicates the number of equivalent line width that can be observed within the reduced image size W_{crop}.

$$K_{hor}\left(W_{crop}\right) = \left\{W_{crop}/\left(w'_{chart/hor}/9\right)\right\} \tag{14}$$

Now that we are working with an image size equal to W_{crop}, the nominal spatial frequency requirement will be given by Eq. (6) or (7). And the limit resolution (MTF) will come at the position where the multiplication factor $P_{hor}(W_{crop})$ is calculate by Eq. (15) as following

$$P_{hor}\left(W_{crop}\right) = \left\{MTF10_{MIN(1:1)/hor}\left(W_{crop}\right)/K_{hor}\left(W_{crop}\right)\right\} \tag{15}$$

and obviously, this value turns to be equal to the value calculated from Eq. (38) given in Sect. 7.5.3.1 of ISO 16505:2015 because both the numerator and the denominator shall compensate for the evaluation image size reduced from $W_{monitor/hor}$ to W_{crop}, cancelling each other change; and furthermore, it is just the evaluation window size that was modified and the appearance of chart will not change by the difference in evaluation window. A minor deviation may occur when using Eq. (5) due to increased deviation of the assumption used in the Eq. (4).

To know the spatial frequency of the chart at a certain position, just multiply the reference spatial frequency at position 1 by multiplication factor indicated by its position. Now, we know how to find a specific point on the hyperbolic chart where the required minimum limit resolution (MTF) point may come or how to reversely read the spatial frequency from a specific point on the hyperbolic chart.

Now, the position on the chart determining the spatial frequency of the required limit resolution (MTF) is known. If the first point in which the periodicity of the chart gets lost by observing the chart output image on the CMS monitor from lower frequency to higher frequency comes at higher frequency to above obtained limit resolution (MTF) point, then it is likely that the CMS could satisfy the requirement. To confirm, evaluate the modulation at that specific point and verify that the modulation is over the 10 % acceptance level. A conventional optical system is generally a decreasing function relative to the spatial frequency and it automatically means that the MTF10 point will come at higher frequency to the above point.

8.4 Preparation and Use of the Parallel Line Chart

The limitation of resolution at high frequency may occur due to the optical limit as well as the design limit caused by limited pixel counts or further combination with image processing.

In case the performance of the CMS is limited by the occurrence of abnormality of the periodic signal causing difficulty to correctly determine a point where the modulation drops to 10 %, an alternative method using a parallel bar chart shall be applied to verify the compliance of the CMS. In this case, a parallel chart containing black and white lines with a spatial frequency equivalent to the required resolution (MTF) shall be prepared and the distinguishability of the line in the reproduced image on the CMS monitor shall be verified.

In principle, the simplest way to prepare a parallel bar chart for verification to compliance is by creating line width and spacing such that it is equal to arc length formed by the viewing angular size in formula (1) as this is the minimum detail that is required to be distinguishable. It should be compose by several line separated by same space to create a periodic pattern which is necessary for the visual evaluation. The result is near equivalent to the spatial frequency in the real world plane such that a width of the line and separating space is obtained by dividing the real world width formed by the viewing angle $\alpha_{monitor/hor}$ by $MTF10_{(1:1)hor}(W_{monitor/hor})$ because when captured by the CMS, it should display $MTF10_{(1:1)hor}(W_{monitor/hor})$ number of line width along the monitor defined size. Other image size like $W_{monitor/hor/min}$ or W_{crop} can also be used. When using the required minimum field of view, the single line width is obtained by dividing the width of real world plane by the $MTF10_{(1:1)hor}(W_{monitor/hor/min})$ and when using a cropped image, divide the width of the real world plane by $MTF10_{(1:1)hor}(W_{crop})$. Otherwise, if determining the chart pattern in the same procedure to derive the Eq. (11) or the equivalent Eq. (38) of ISO 16505:2015, the spatial frequency at the position determined by the calculation shall be duplicated as parallel bar chart for the limit resolution verification. Whichever calculation procedure is used, all results in principle contain the same physical spatial frequency.

Prepare a printed version with 5 black lines (at least 4 according to ISO 16505:2015) with the calculated width separated by a white space of the same width. Some printer may exhibit augmented tone at the black line edge resulting in unsuitable non-uniform tone. The entire chart shall contain a uniform tone avoiding any contour processing of the chart.

The parallel chart shall be oriented with a rotational angle of about 5° to avoid large aliasing on particular lines of the chart. The distance to the chart shall be kept according to the original distance used to calculate the spatial frequency of the chart. Note that by moving the chart to closer distances, the spatial frequency of this parallel bar chart will decrease, and vice versa. Exaggerated rotation θ of the chart shall be avoided as the rotation of the chart causes the line width of the chart to be increased by 1/cos(θ) in targeted horizontal/vertical orientation and the spatial frequency to be decreased by $\cos(\theta)$. When placing a chart with 5° rotation this deviation is considered to be negligible.

Verification to the compliance shall be made by comparing the output images captured on the CMS monitor and compared to the Guide of visual acceptance level provided in the Fig. D.2, Annex D.3 of ISO 16505:2015.

8.5 Using SFR Method for Measuring the Resolution on Image Displayed on the Monitor

Spatial Frequency Response method, abbreviated as SFR, is becoming popular as a tool for evaluating image resolution of electronic still picture systems. Details of the technology and measurement are provided in ISO 12233:2014. And related information is also available on the links provide by the Society for Imaging Science and Technology at (http://www.imaging.org/ist/resources/standards/Digital_Camera_Resolution_Tools.cfm) [3].

SFR method is useful for analytical evaluation when used in a system exhibiting a linear characteristic where input optical signal are linearly processed to the output signal. But it might be largely influenced by the non-linear image processing and not suitable in some case if improperly used.

A CMS system is likely to take advantages of state-of-the-art new technologies making use of image processing to enhance the visualization of the captured scene. Many of such image processes modify the edge characteristic of the captured image, and each product exhibiting specific characteristic particular to each product design, and further exhibiting different response characteristics according to luminance condition of the scene. Furthermore, most of such CMSs are expected to operate in a fully automatic mode. Evaluation by disabling all such functions will result in different image and will not reflect the actual use condition performance in the real case.

Therefore, the SFR method alone may not provide a suitable result when evaluating a CMS. Regardless of all these reproducibility problems, ISO 16505:2015 kept the SFR method as a part of the evaluation because the chart used in the evaluation is simple and it is also practical to be used at working distance of a CMS when charts are placed several meters away from the camera. There is always a clear correlation between the SFR curve and a modulation transfer function measured by direct observation of the chart particular to each product and measuring condition. So, ISO 16505:2015 provides an informative annex to create an off-set correlation curve between values obtained from SFR method and modulation transfer function directly evaluated from the hyperbolic chart evaluation method, and further incorporating the visual limiting resolution point onto the correlation curve.

The SFR method described under ISO 12233:2014 requires the evaluation of luminance intensity to be worked in the linear domain by inverting the opto-electronic conversion function according to ISO 14524. ISO 16505:2015 has not made this requirement mandatory because of the difficulties to make an analytical approach of the integral evaluation of CMS.

The SFR method provide a SFR(*f*) curve as a function of spatial frequency and a single slanted edge measurement using SFR method may provide both the MTF10 and MTF50 spatial frequency. The result of SFR method could be affected by the illumination condition or chart contrast. Changing the evaluation condition of different measurement like the depth of field measurement shall be kept to a minimum and if the evaluation conditions are changed, the correlation shall be confirmed accordingly to avoid deviation in the results. Due to the difficulties to illuminate the entire field of view at distance, some hints are provided in ISO 16505:2015, for example, using a small foreground chart to cover the rest of the field of view around the measuring single hyperbolic chart.

The informative Annex E in ISO 16505:2015 provides a guide how to create a correction curve of the off-set deviation according to spatial frequency. The reproducibility of such a look-up table is likely to be very vulnerable to measuring ambient light condition and also on the chart contrast condition as image processing is applied to tonal edge creating the non-linearity. This handbook will not go into any further detail of the SFR methods because there is not any universal way to remedy and avoid all related side effects when the CMS is evaluated, including effects of the output device performance.

The SFR method analyses the captured image from the discrete sampled pixel data and estimates the output response curve plotted according to the spatial frequency "normalized" by the discrete sampling frequency. When evaluating a system which is sampled by N pixel in the orientation of interest, the normalized spatial frequency is equal to 1 for these N pixels sampling. Therefore, to convert a spatial frequency given in normalized spatial frequency to a unit in [LW/PH] considering a square image, multiply the normalized spatial frequency values obtained by SFR method by the number of pixels N composing the width in the orientation of interest.

For the integral evaluation of a CMS, the output image displayed on a monitor is captured using a reference camera. The SFR sampling frequency is dependent on the reference camera alone and independent of the CMS. Assuming that a square evaluation area is captured and cropped by the reference camera with pixel sampling count of $N \times N$ pixel, the sampling frequency is equal to 1 for signals containing N sampled data equivalent to N row in height, which gives $N \times 1$ [LW/PH]. Details of unit conversion for SFR measurement is found in Annex C of ISO 12233:2014.

The pixel count of the reference camera can be determined independently of displayed image on the CMS monitor. Nevertheless, an oversampling at least by a factor of 4 combined with low pass filter is always recommended to minimize aliasing effects of the reference camera.

The normalized spatial frequency is dependent on the measuring device pixel numbers while the spatial frequency given in [LW/PH] units is independent to the measuring pixel and only dependent on the image height. Therefore, spatial frequency value converted to [LW/PH] in the reference camera remains valid within the image displayed on the monitor, whose pixel counts is independently determined from the reference camera pixel count.

Example on how to work with the unit:

Following is an example of the SFR measurement and related units. Consider evaluating a partial image on the CMS monitor in which the required field of view on monitor is displayed in a $W_{monitor/hor/min}$ = 90 (mm) and an area size of $W_{crop} \times H_{crop}$ 40 (mm) × 40 (mm) on the monitor panel is capture and cropped with a reference camera. And consider using a reference camera to capture this area in pixel counts of 480 (px) × 480 (px). If the SFR result is presented as normalized spatial frequency in respect to its sampling frequency, then it should be converted to [LW/PH] units. A normalized spatial frequency value of 0.2 [sampling cycles per pixel], for example, is equivalent to 480 × 0.2 = 96 [LW/PH]. If this value is a nominal sharpness value as measured in this cropped image $MTF50_{(1:1)/hor}(W_{crop})$ = 96 [LW/PH], the nominal sharpness in a square image $W_{monitor/min}$ can be reversely calculated using Eq. (8).

$$\begin{aligned} MTF50_{(1:1)/hor}\left(W_{monitor/hor/min}\right) &= \left(W_{monitor/hor/min}/W_{crop}\right) \times MTF50_{(1:1)}\left(W_{crop}\right) \\ &= (90/40) \times 96\,[LW/PH] = 216\,[LW/PH] \end{aligned}$$

As we can see from this example, the measurement of the partial image can be handled independently of the pixel counts in the original image on the CMS monitor. The result of the evaluation done within the reference captured image given in [LW/PH] unit is evaluated in a self-consistent manner, and the same value is also valid on the respective window in the CMS monitor, and the nominal spatial frequency is reversely calculated using the Eq. (8) or (9), according to the definition of the monitor size.

9 Sharpness

Besides the resolution (MTF), the sharpness plays an important role in the perception of the environment. Our environmental perception behaviour is composed by multiple sequences before we observe the target scene with our central vision where we see a target with best acuity, recognizing detail of an object. Prior to this precise observation of the particular small area of the scene by our central vision, we perceive the likeliness of an object by the so called “rod” which composes most of our visual sensitivity in the periphery of the fovea at the centre of our vision. These rods have a lower visual acuity but are highly sensitive to motion with stimulus generated by the temporal change of the light achieving these rods, which is a result when edge of an image moves across these rods. A high spatial frequency image is well visible by our central vision even at relatively lower modulation amplitude but these rods on the other hand require higher modulation amplitude so that they become perceived at the periphery of the central vision.

A comprehensive perception and interpretation of the environment is achieved by a combination of the detailed central vision information combined with the peripheral visual information both playing an important role in the fast interpretation of what is occurring around.

Apart from resolution (MTF), a second parameter is adopted to characterize the performance of CMS considering above needs of higher modulation amplitude. And the spatial frequency at the point where the modulation transfer function signal drops to 50 % is the parameter adopted to represent such requirement as “sharpness” characteristic of the CMS, and it is denoted as MTF50.

In brief, sharpness helps in recognizing content likeliness of a scene in a short glimpse of view without gazing into details of the scene while the resolution (MTF) refers to an indication factor that reflects the capability to see the limiting higher frequency details of the scene by our central vision.

Testing procedure to measure the modulation can be performed in a similar manner as for verifying the limit resolution (MTF). The required spatial frequency is half of the requirement for resolution (MTF). As a result of the lower spatial frequency compared to the evaluation of the resolution (MTF) requirement, a system satisfying the resolution requirement is likely to exhibit less artefacts in this lower frequency range which leads to the loss of correct periodicity of the test chart. The periodic output signal of the test chart is likely to be correctly observed without the need of using a separate auxiliary parallel bar chart in the sharpness requirement frequency range if resolution (MTF) requirement is properly met.

With an appropriate low pass filter, the moire effect of the sampling or monitor pixel design will not be apparently affecting the observation of the modulation transfer function at this low frequency range of sharpness requirement. The measurement is the same as the measurement for resolution but the spatial frequency is obtained from the direct observation of the point on the reproduced image of the chart where modulation decreases down to 50 % of the reference black and white level, instead of the 10 % for the limit resolution. Or optionally one can indirectly judge that a CMS whose modulation is over 50 % at the required minimum spatial frequency for sharpness (see right side of Eq. (33) and (34) of ISO 16505:2015) complies with the sharpness requirement. Considering that the MTF(f) curve is generally a decreasing function of spatial frequency, the above verification indirectly secures that the spatial frequency where the MTF(f) drops to 50 % is over the above required minimum spatial frequency as sharpness minimum requirement.

Besides the resolution (MTF) and sharpness, contrast also plays an important role in the visual perception. The modulation transfer function itself is just a relative capability and is a normalized value in respect to the signal at zero frequency modulation. Contrast is handled as a separate characterization parameter of the CMS in Sect. 6.7.2 of ISO 16505:2015.

10 Others

The CMS is expected to be evaluated in an automatic exposure operation mode. Therefore, incorrectly positioned or incorrectly illuminated charts may result in underexposure or overexposure of the CMS camera. An overexposure image may result in erosion of black region by the excess light while an underexposure image may result in considerable degradation of the image by the dark level noise. Both conditions may affect the input and output characteristics of the CMS and making it difficult to obtain the typical MTF characteristics of the CMS. To avoid inappropriate exposure adjustment of the CMS by the automatic exposure operation during the evaluation, the entire field of view of the CMS camera should be covered by a near 18 % reflectance neutral colour grey background, illuminated in same condition as the main area of the test chart.

The illumination condition of the chart is not directly mentioned as part of the measurement procedure and condition described for the resolution requirement in ISO 16505:2015 but the requirement given under the same standard for sharpness evaluation states to illuminate the chart uniformly by a diffused light source of type D65 and same condition shall be applied.

ISO 12233:2014 adopted a lower contrast edge chart making the method less sensitive to the non-linear image processing. This helps in reducing the effects of image processing like edge enhancement but it cannot yet provide a stable results as the non-linear image processing is design and product dependent. All measurement suggested under ISO 16505:2015 was based on black and white chart but use of low contrast chart may also help to minimize the appearance of false signal when evaluating the modulation transfer function.

There are several factors which influence the resolution of the CMS. The number of pixel which consist a CMS shall at least be capable to sample and display the captured scene image. In general, a sampling frequency higher by at least 1.5 times of the target minimum spatial frequency is preferred to cover for capturing a natural scene. Several other factors like photodetector fill factor or detailed layout of the photodetector array or post-filtering of those detected arrayed image will affect the results of the limit resolution. For example, in a system targeted to distinguish at least 300 [LW/PH], a selection of components of at least $1.5 \times 300 = 450$ sampling pixels in the target orientation might be preferable to avoid considerable aliasing. The suggested value of sampling factor 1.5 times higher to the target spatial frequency is a just reference factor but it does not necessarily secure the reproducibility of the system to distinguish whatever signal is.

Note: There is an erratum in the description of the symbols $\mathrm{MTF10}_{\mathrm{MIN(1:1)/hor}}$ and $\mathrm{MTF10}_{\mathrm{MIN(1:1)/ver}}$ in Sect. 4, Symbols and abbreviated terms of ISO

16505:2015. It is a definition of the horizontal/vertical spatial frequency requirement of resolution (MTF); whose nominal value is dependent on the evaluation window size.

For a better tracking of the nominal value, this handbook suggests working with the above parameters with the denotations of the image size shown as arguments in parenthesis.

11 Conclusions

The resolution (MTF) and sharpness measurement in a CMS has to be evaluated as an integral performance of the whole CMS system. And the requirements are defined according to the orientation of interest and not necessarily referencing to the actual full image height as is the case of ISO 12233:2014 definition. Therefore, when the [LW/PH] unit is used, there are several aspects in which the definition had to be modified compared to the resolution definition under ISO 12233:2014.

The denotation of the nominal spatial frequency in this Chapter takes the image window size as a variable to explicitly show that the value correspond to that specific image size under evaluation.

The requirement defined under ISO 16505:2015 takes into consideration the minimum level of the scene distinguishability according to the lower limit determined by the driver visual acuity allowance. The requirement defined as minimum requirement may not provide equivalent visual distinguishability to those driver with higher visual acuity driving with traditional mirror. Hence, it is always preferable to target a higher performance compared to the target given under ISO 16505:2015. The author will be glad if some of the information provided herein serves as a hint to improve and provide a further safer CMS product to the automotive market and driver's expectation, exploiting the CMS capability beyond a traditional mirror.

Each regulation and/or target user has specific target for visual acuity or convex mirror allowance and CMS shall be developed considering each particular target market and usage condition.

References

1. ISO 16505:2015, Road vehicles—Ergonomic and performance aspects of Camera Monitor Systems—Requirements and test procedures
2. ISO 12233:2014, Photography—Electronic still picture imaging—Resolution and spatial frequency responses
3. "Digital Capture Resolution Measurements" provided by IS&T, the Society for Imaging Science and Technology. http://www.imaging.org/ist/resources/standards/Digital_Camera_Resolution_Tools.cfm. Accessed 28 Sept 2015

Vision in Commercial Vehicles with Respect to Camera Monitor Systems

Patrik Blomdahl

Abstract Commercial vehicles show big improvement potentials for drivers when rear-view mirrors are replaced by Camera Monitor Systems. At the same time there is a wide range of specific requirements that have to be considered to reach successful integrations. This is because many use cases are unique and originate from how and when these large and often complex vehicles are utilized to transport goods or people. The present text summons information about a wide range of aspects that need to be considered when CMS is introduced in commercial vehicles. Important situations related to both driving and non-driving are described. It is explained how requirements should be established to secure that CMS will provide the same or a higher level of performance compared to traditional mirrors. Specific focus is put on needs for changed fields of view. When cameras and monitors are installed on the vehicles, it has to be remembered that drivers, as human beings, should still be able to cope with and understand the new system in a natural way. CMS can provide several benefits over the traditional mirrors, but there are nevertheless aspects that need to be handled differently. If they are considered, there is a big potential to improve driving situations significantly regarding both indirect and direct vision. In the future, developments can start from a new technical foundation which will facilitate a multitude of additional opportunities that could influence the driver environment, the driving task and the transport system, as we know it.

Keywords Commercial vehicles · Indirect vision devices · Vehicle combinations · Traffic safety · Intelligent transports

---

P. Blomdahl (✉)
Vehicle Engineering, Volvo Group Trucks Technology, 405 08 Gothenburg, Sweden
e-mail: patrik.blomdahl@volvo.com

A. Terzis (ed.), *Handbook of Camera Monitor Systems*,
Augmented Vision and Reality 5, DOI 10.1007/978-3-319-29611-1_4

1 Introduction

Commercial vehicles and the way they are driven are in many ways very different from passenger cars. This chapter describes these differences, and some are particularly important to remember when future commercial vehicles will be developed with Camera Monitor Systems instead of traditional rear-view mirrors.

The first obvious thing that most people notice is that commercial vehicles are bigger and take up more space in the traffic environment. The size can often come from their main purpose of transporting various goods or people. In the case of goods, the total weight and/or volume that can be carried get important. In the case of people, the numbers of persons that can find seated, standing or lying room become key.

Due to these needs to optimize the vehicle to the amount of goods or number of people that can be transported, the layout of the commercial vehicle has become different than the layout of the passenger car. It is common that the driver is positioned higher above the ground and further in the front of the vehicle. Besides, the need to prioritize how much can be transported and how goods or people can be handled in efficient and safe ways leads to smaller areas of windows that the driver can use to overview the surroundings. Such considerations, together with sheer vehicle size and additional complexity that can come from additional trailers, means that indirect vision via mirrors or camera systems become essential additions to the direct vision via window openings.

In the transport system, commercial vehicles are in general used to pick up goods or people at almost any hour, day and night. To support this need for transport, there are in many cases specifically arranged terminals and in some cases even specific lanes or roads that the commercial vehicles can use. In most cases, however, commercial transports must be made within the infrastructure used by all other road-users. Solutions for the vehicles and the infrastructure should be chosen to secure that all road-users can observe and keep track of each other in a good way. Besides, it is important to consider and support good communication and understanding between the different road-users.

There are some inevitable compromises when traditional mirrors, obeying to the laws of physics, are placed onto the adapted layouts of commercial vehicles. Mirrors cannot always be placed where they are easiest to overview. In order to cover sufficient fields of view without getting too big, which would result in unacceptable blind spots, convex mirrors with smaller glass radii are allowed. However, in their turn, they lead to more distorted views, which make it harder for drivers to understand what is being seen.

All in all, this means that replacing traditional mirrors with Camera Monitor Systems has the potential to greatly improve vision conditions for the drivers of commercial vehicles. This chapter intends to bring light onto some aspects that have to be observed to successfully utilize this potential improvement.

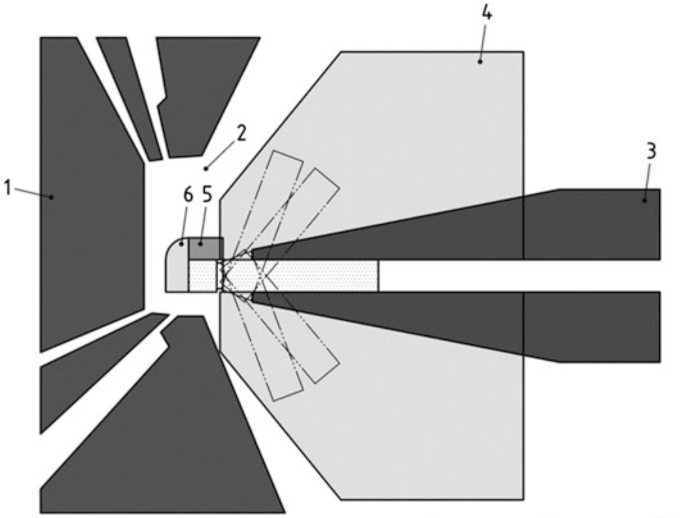
2 Relation Between Direct and Indirect Vision in Commercial Vehicles

Figure 1 helps to illustrate the specific vision considerations above [1]. In comparison with passenger cars, the size of typical blind spots as well as the space taken up by the commercial vehicles can come as a surprise [2]. The figure shows a top view of which vision areas are typically provided by different means of direct and indirect vision. The vehicle combination is the very common truck tractor with connected semi-trailer that folds as the vehicle turn around a corner.

For a passenger car, both the blind-spots and the areas covered by indirect vision are significantly smaller. This is illustrated in Fig. 2.

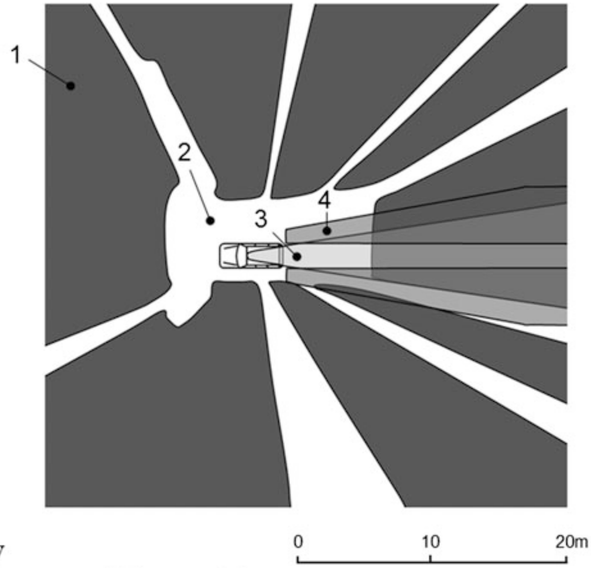
An investigation has been made where truck drivers were asked about the importance of different means of vision [3]. In total 37 drivers participated.

First the drivers provided indirect vision ratings regarding the importance of the views provided by the different types of traditional mirrors. The rating scale ranges from 1 to 5, where 1 corresponds to a mirror used most often and 5 corresponds to a mirror used most rarely.



Key
1 Areas of direct vision
2 Areas of blind spots
3 Regulated field of class II main rear-view mirrors
4 Regulated field of class IV wide-angle rear-view mirrors
5 Regulated field of class V kerb-view mirror or camera
6 Regulated field of class VI close-up front mirror or camera

Fig. 1 Ground area covered by different means of vision [1]



Key
1 Areas of direct vision
2 Areas of blind spots
3 Regulated field of class I interior rear-view mirror
4 Regulated field of class III main rear-view mirrors

Fig. 2 Relation between direct and indirect vision in commercial vehicles and passenger cars

In addition to this, the drivers provided direct vision ratings for the importance of different portions of the window openings. The same rating scale was also used for this. The following figures illustrate the outcomes.

The results in Fig. 3 indicate that the main mirrors are more important in long-haul trucks than in distribution trucks, but it also seems that the passenger side wide-angle mirror has additional importance in distribution transport.

Importance: 1= used most often 5 = used most rarely	Left main rear-view mirror	Right main rear-view mirror	Left wide-angle mirror	Right wide-angle mirror
Long-haul truck drivers, average:	1,00	1,24	2,06	2,24
Distribution truck drivers, average:	1,71	1,71	2,29	1,43
Total average:	1,20	1,32	2,29	2,60

Fig. 3 Importance of each rear-view mirror based on how often they are used

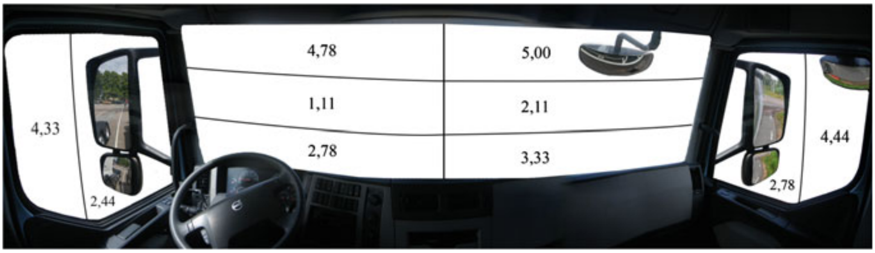


Fig. 4 Importance of different window areas in the case of a long haul truck (1 means used most often, 5 means used most rarely)

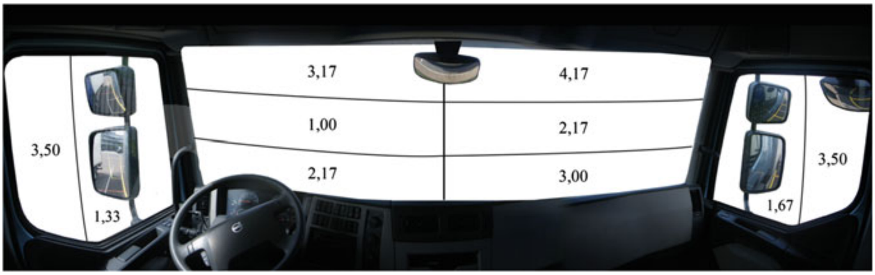


Fig. 5 Importance of different window areas in the case of a distribution truck (1 means used most often, 5 means used most rarely)

The results in Figs. 4 and 5 indicate that the direct vision through the front portion of the side windows as well as over the dashboard in front of the driver is significantly more important in distribution trucks than in long-haul trucks.

There is also a combined usage of direct and indirect vision that is to be remembered. This combination should be considered when positioning indirect vision views in relation to the direct vision provided through different window openings [4]. Depending on use case, the reference direction of the direct vision can be either forwards through the windscreen or sideways through the side windows as is indicated in Fig. 6.

A consideration opposite to the above occurs when driving at night. Then a bright display must not cause disturbing glare [2, 4, 5]. Hence, in these cases a display that is brighter than the surroundings should not be placed too close to the reference direction where key direct vision information needs to be observed.

In the same above investigation [3], information has also been gathered to consider which the main usages for each rear-view mirror are, and what the drivers would like to see in them. Figures 7 and 8 show this type of information for the left and right side rear-view mirrors.

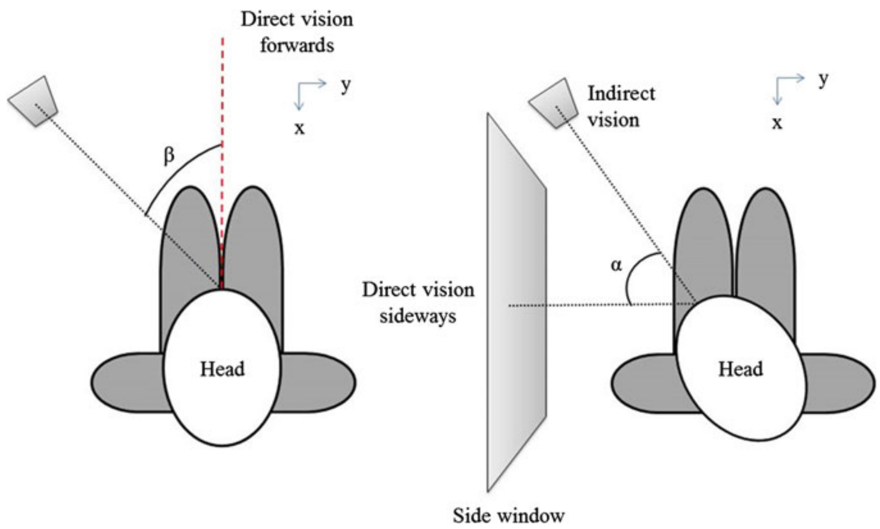


Fig. 6 Relation to reference directions where direct vision is provided [4]

Left main mirror

- "Use it for reversing"
- "Use it frequently to check the traffic behind"
- "Want to see 30 % of the truck and 70 % of the surrounding environment"
- "Want to see from the rear part of the cab, along the side of the carriage, and the whole lane to the left to the truck"

Left wide-angle mirror

- "Use it to see the whole carriage when the main mirror does not cover it"
- "I rather open the door than trust the wide-angle mirror when reversing (difficult to judge distance due to distortion)."

Fig. 7 Rear-view mirror usages on the *left side* [3]

3 Use Cases for Commercial Vehicles

3.1 General Categories of Use Cases

In consideration of the total list of use cases within ISO 16505 [1] and the above main reasons for vision considerations, the following general categories of use cases need extra focus for commercial vehicles:

Fig. 8 Rear-view mirror usages on the *right side* [3]

A. Adjacent traffic in lanes and connecting roads
B. Relation to surrounding infrastructure with varying sizes of the driver's own vehicle
C. Light and weather conditions in relation to vehicle size and how cameras and monitors are installed
D. Driving in complex traffic with many vulnerable road users
E. Longer periods of driving in darkness
F. Usage when and after living in the vehicle

3.2 *Driving-Related Use Cases*

To illustrate the use case categories related to driving, the criticality of different traffic situations and driving conditions is used. This input also originates in the mentioned investigation [3] that has been made with truck drivers, see Fig. 9.

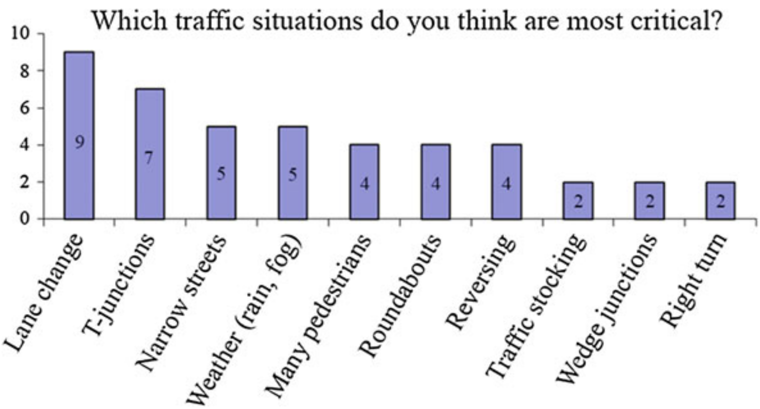


Fig. 9 Traffic situations mentioned as critical [3]

Fig. 10 Lane change condition

Among all use cases, the following are of specific importance regarding the fields of view provided by the indirect vision devices [1]:

(1) Lane changes where other road users must be observed
(2) Turning left or right around corners with connected trailer
(3) Maneuvering towards an exact location for leaving or picking up people and/or goods
(4) Driving through a roundabout with trailer connected
(5) Pulling out into a main road from an attaching road, at an angle

These have been identified since they need additional head movements with traditional mirrors if truck drivers should have sufficiently large fields of view. The results come from a pair of independent investigations [3, 4] with a total of 24 experienced truck drivers driving in various types of traffic conditions, but similar results are expected for other commercial vehicles.

The following Figs. 10, 11, 12, 13 and 14 illustrate use cases where fields of view often become an issue.

When turning around corners in city areas, the problem to see a sufficiently large field of view within the indirect vision device gets combined with the problem to follow a rapidly changing scenario that includes dense traffic and vulnerable road users. Solutions to counteract these problems are increased areas covered by the close-up vision devices as well as additional sensors that help to notify drivers about adjacent objects.

The semi-trailer in itself causes blind spots when the vehicle combination gets folded. However, the biggest problem in a right turn as illustrated is that it becomes hard for the driver to see the semi-trailer rear end on the passenger side. It is impossible in the main mirror so the wide-angle mirror has to be used, but it provides a distorted view where details are hard to grasp.

Fig. 11 Turning left or right around corners [3]

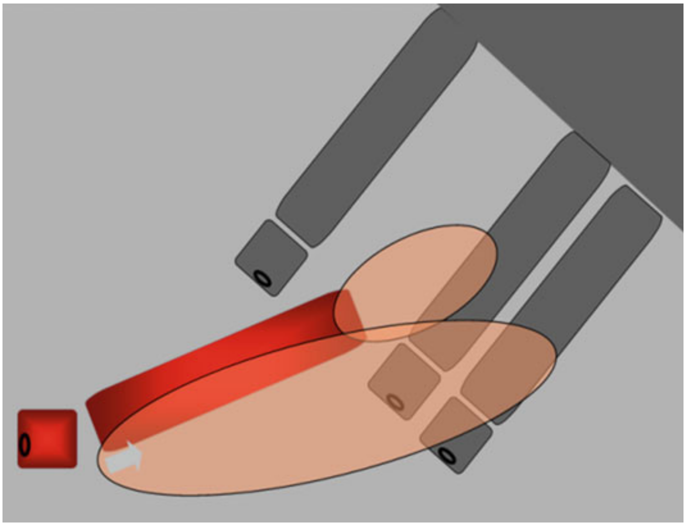


Fig. 12 Maneuvering towards an exact location to leave or pick up people and/or goods

Figure 13 shows a situation with a tractor truck and a semi-trailer, but the same condition applies to e.g. a passenger car and a caravan. By moving the head the driver can see the rear end of the semi-trailer in the mirror.

Wedge junctions of different kinds often cause additional problems for big commercial vehicles as the driver will need to position the vehicle in the best possible way to facilitate the turn.

Fig. 13 Driving in a roundabout with a trailer connected [3]

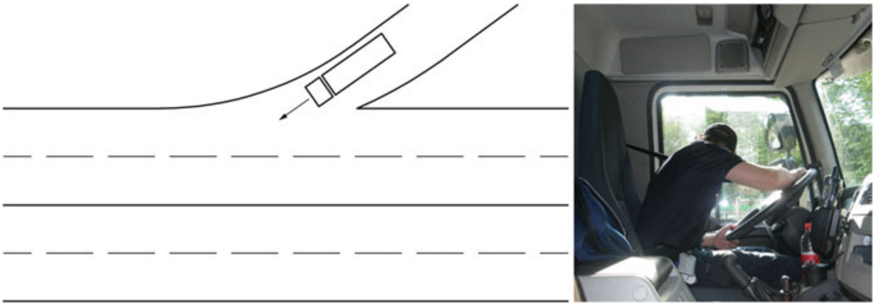


Fig. 14 Pulling out into a main road from an attaching road at an angle [1, 3]

3.3 Use Cases not Related to Driving

Indirect vision devices, mirrors or cameras, are not only used while driving. There are also use cases where indirect vision is required at standstill. Think for example about the most common situation when a passenger car has stopped at the roadside, and the driver wants to look out for approaching traffic from behind before leaving the vehicle.

In the case of commercial vehicles, use cases related to standstill situations are often slightly different from those involving passenger cars or motorbikes. The main reasons for this difference are: drivers and potential passengers can live in the vehicle for longer periods of time; it is more common to stay and relax in the vehicle (e.g. while waiting for the turn to unload goods); and the vehicle stands still with the engine on idle while the driver moves in, out and around the vehicle, e.g. when loading and unloading goods of different kinds.

The common practice of living or relaxing in the vehicle originates from the driving time regulations in many countries. These regulations specify that drivers must have a break with a proper time period of rest after a certain number of driving hours. As traffic is unpredictable and transport takes place also in remote areas, the drivers and potential passengers will therefore need to stay in the vehicle whenever it is time for a rest.

The Camera Monitor System is only one among many other mechanical or electronic sub-systems which are integrated into the complete vehicle system. Therefore it is important to study each driver and passenger usage in more detail to secure that the CMS can support the users in an optimal way without causing annoyance, like unnecessary delays, or risk for surprises, like unexpected behaviour.

When designing the system according to the specific use cases, it is beneficial to keep track of the optional triggers that can be used to activate the Camera Monitor System and of the typical time the follow periods that it takes for the driver to perform each step included in the use case.

Below follows five typical use cases at standstill which involve utilizing the indirect vision systems, mirrors or cameras, in ways that are specific for commercial vehicles [1]. To simplify, drivers or passengers are sometimes called occupants.

3.3.1 Ordinary Entry in Order to Take Off

This use case starts with the driver unlocking the door from the outside. After the door is opened, the driver enters the vehicle, gets seated and applies the seat belt. As soon as the ignition is activated, the driver can check if traffic approaches from behind by using the indirect vision device. In comparison with the driver of a passenger car, the driver of a commercial vehicle cannot, to the same extent, rely on direct vision through windows as an alternative to the indirect vision device. If the road is clear, the parking brake is released and the vehicle can take off into the road.

3.3.2 Ordinary Exit After Coming to a Halt

At the start of this use case, the driver brakes the vehicle until it stops at the road side. When the vehicle is at a standstill, the parking brake is applied. If the driver intends to leave the vehicle, the ignition is normally turned off and the seat belt is removed. After that, any traffic that approaches from behind has to be checked via the indirect vision device. If the coast is clear, the driver opens the door from the inside and leaves the seat and the vehicle. The door is closed from the outside and potentially locked in order to end the use case. The same use case is also valid for a passenger who leaves the vehicle, but in that case the steps will not include the vehicle interactions of applying parking brake and handling the door locks, as these are considered as tasks performed by the driver.

3.3.3 Getting Out After a Longer Time Period at Standstill

This use case appears in vehicles where drivers and/or passengers stay in the vehicle for a longer period of time in order to rest, wait or sleep which can even be overnight.

The occupants have been inside the vehicle for a longer period of time. It is then also likely that the vehicle is locked up from the inside. When the driver and/or passenger want to get out, the first action will in that case be to unlock the door from the inside. After that, the occupants need to check if there is any traffic around the vehicle including vehicles that might approach from behind. Here it becomes of utmost importance that the indirect vision device can provide a rearwards view as it is often impossible to notice other road users by opening the door and very inconvenient to bend out through side windows that would also first have to be opened.

When it is considered possible and safe, any door is opened from the inside and the occupant leaves the vehicle which does not necessarily need to be via sliding across the driver or passenger seats—which could otherwise have formed an opportunity for trigging the system via pressure sensors. Once on the ground, the occupant shuts any door from the outside and probably also locks the vehicle before walking away from the parked vehicle.

3.3.4 Taking off Directly After a Longer Time Period at Standstill

The application of this use case commonly occurs when the driver has to continue the journey after having rested, slept or stayed in the vehicle for a longer period of time. After getting ready for a new period of driving—that might not necessarily involve leaving the vehicle as part of any kind of morning routine—the driver will get seated. Then the ordinary start up procedure will follow according to the sequence of putting on the seat belt, activating the ignition, and releasing the parking brake. Before taking off into the traffic, it is necessary for the driver to check that no obstacles are in the way. This includes checking if there is any traffic approaching from behind. Here the indirect vision devices need to be fully active to support a safe start of the driving.

3.3.5 Getting Out of the Vehicle While the Engine Is Still Running

For commercial vehicles, it is quite common that the driver leaves the vehicle with the engine running while he or she takes care of goods, passengers or makes some necessary arrangements with the surroundings. Examples of common driver activities in these cases are that gates into confined areas have to be opened or that bodybuilder equipment has to be used to handle the specific type of goods. The tailored bodybuilder equipment will then in most cases need power from the vehicle that is facilitated by the running engine.

A bodybuilder company adds a specific superstructure or body to a truck in order to adapt it to the type of goods it will carry. Figure 15 provides some examples of bodies added by bodybuilders.

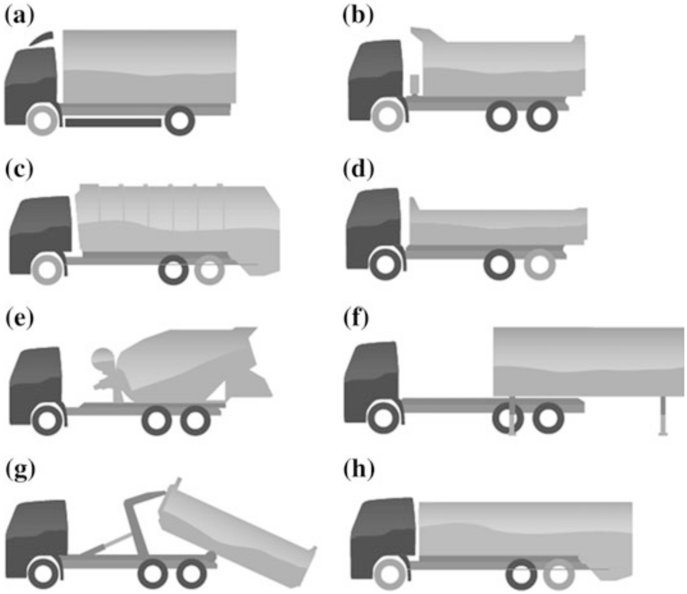


Fig. 15 Examples of rigid truck superstructures. **a** Van body. **b** Dump body. **c** Refuse body. **d** Tipper. **e** Concrete mixer. **f** Swap body carrier. **g** Hook lift. **h** Tanker

Under such circumstances, the vehicle will be at a standstill with idling engine, the gear in neutral, and parking brake applied. If bodybuilder equipment is needed to load/unload goods, the power take off from the vehicle will be activated. The driver will potentially remove any seat belt before opening the door, leaving the seat, and getting out onto the ground. However, before the driver opens the door, it will first be necessary to look out for traffic approaching from behind. In the same way as before, the indirect vision devices become essential to verify that it is safe to leave the vehicle. Once on the ground, the driver will most often shut the door from the outside and sometimes also lock it to prevent any unknown individual from entering in his or her absence.

4 Categories of Commercial Vehicles in Relation to Requirement Fulfilment

The requirements that the CMS should fulfil will come from an analysis of the actual driving situations of a certain vehicle combination used for an intended transportation purpose [1]. For this purpose the relevant use cases within the general categories of use cases that were listed above need to be interpreted.

It is, however, possible to establish these requirements by grouping the vehicles and vehicle combinations into vehicle categories with similar needs for indirect

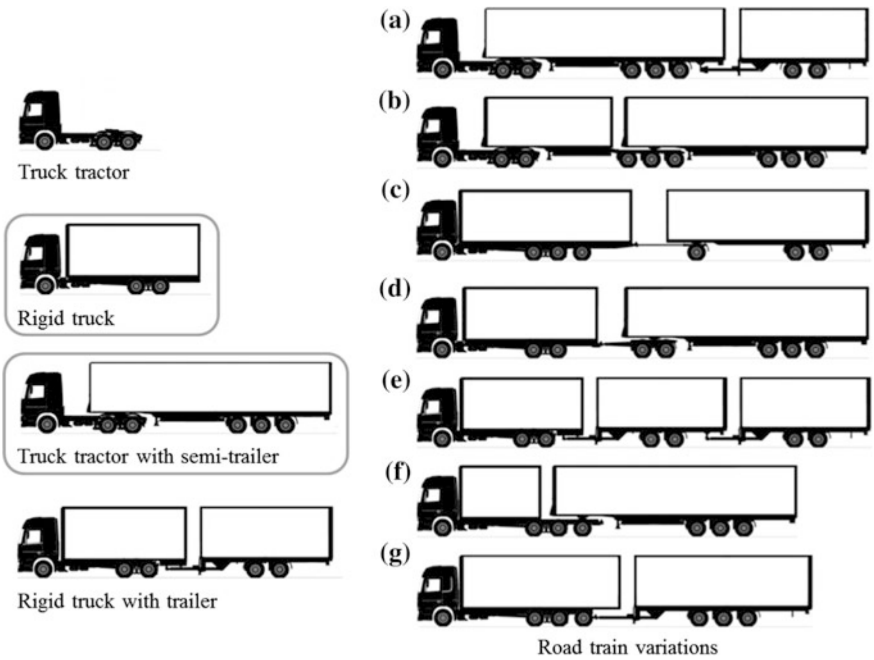


Fig. 16 Typical heavy truck combinations and road train variations with encircled main categories for requirement building [1, 6]

vision [1]. The vehicles in one of those categories will then need CMS that fulfil the same level of requirements. Forming such vehicle categories makes it easier to establish valid requirements and takes away the need to analyse each and every single vehicle variant and combination.

Below in Figs. 16 and 17 are some typical vehicle combinations for heavy trucks and buses respectively.

As will be explained below it is normally sufficient to consider the encircled categories in Figs. 16 and 17 for the definition of requirements to handle changed fields of view.

In the case of heavy trucks, the rigid truck category in Fig. 16 will represent all different usages that rigid trucks have [1]. This category will also represent the most basic case of a truck tractor without any connected semi-trailer. In addition, rigid trucks are equipped with a lot of different superstructures or bodies tailored for each different transport. See Fig. 15 which shows examples of such different superstructures.

The rigid truck at one end represents more basic vehicle configurations; the truck tractor with connected semi-trailer represents all more complex vehicle configurations seen to the right in Fig. 16. This category requires a higher level of CMS requirements regarding covered fields of view [1]. The reasoning behind this is evident from Fig. 18.

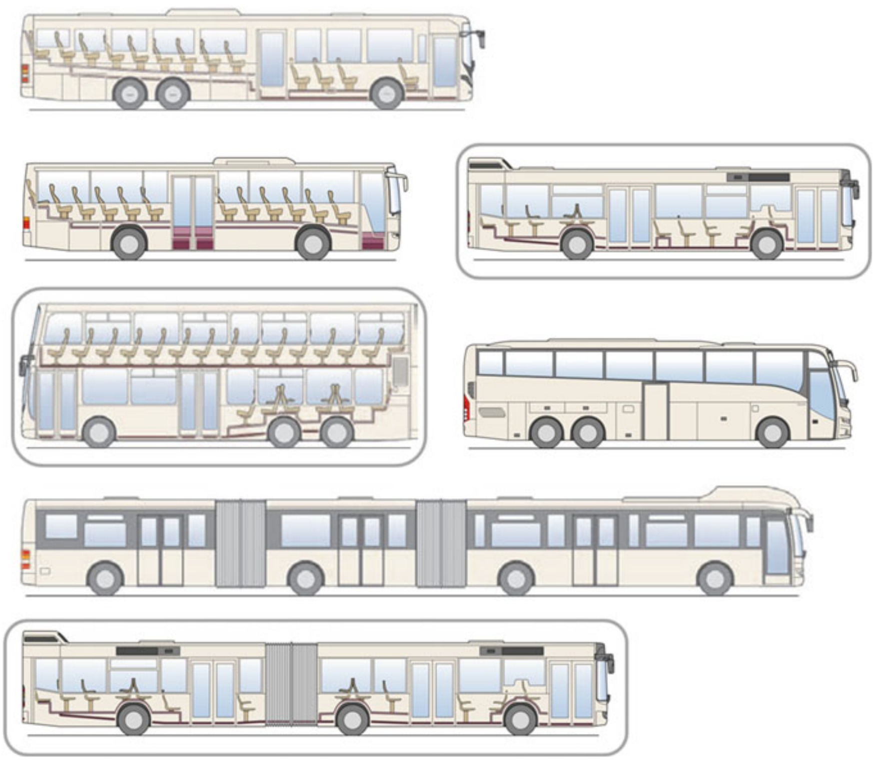


Fig. 17 Typical heavy bus combinations with encircled main categories for requirement building

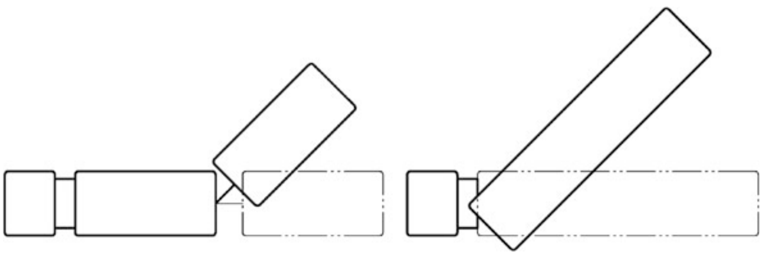


Fig. 18 Illustrated folded situations for two truck combinations—rigid truck with trailer (*left*) and truck tractor with semi-trailer (*right*) [1]

When a rigid truck with connected trailer and a truck tractor with connected semi-trailer are compared, the rear end of the semi-trailer takes up most space on and above the ground. This tells the largest changed fields of view are needed for the latter combination. Hence the requirements will be built for this truck tractor with a semi-trailer combination as a category representing all more complex vehicles.

5 Fields of View—Direct and Indirect Vision

The complexity of the vision for a truck tractor with connected semitrailer was shown in Fig. 1. The area of the ground to be seen is covered by different means of vision: direct vision via window openings as well as indirect vision provided by different classes of indirect vision devices. Figure 1 also provides brief explanations of the different classes. The direct vision to the rear is often blocked by parts of the cab or by the superstructure carrying goods or people. Therefore indirect vision devices constitute the only means to provide the necessary fields of view.

For main Class II and wide-angle Class IV mirrors the fields of view in part cover the same areas of the ground [1]. Nevertheless, the two mirror classes provide different levels of magnification and distortion due to the difference in allowed mirror glass radii. Because of this they can be said to fulfil different purposes. In general, the main Class II mirrors with their larger radii provide better images for observing details whereas the wide-angle Class IV mirrors with their smaller radii provide opportunities to notice objects entering the scene.

In the case of a vehicle combination that folds, as illustrated in Fig. 18, a need for changed fields of view will appear if the folding angle goes beyond a certain value. With traditional mirrors, the driver achieves the changed fields of view by moving the head while viewing the main Class II mirrors [1, 7]. This behaviour adds more detailed information to the larger but less detailed field of view that the wide-angle Class IV mirrors provide. This principle is illustrated in Fig. 19.

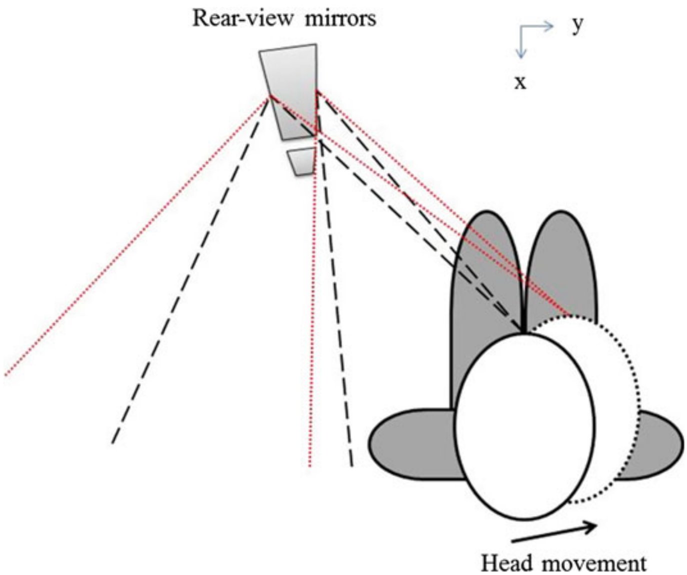


Fig. 19 Panning based on head movements with traditional mirrors [4]

5.1 Fields of View Provided by Indirect Vision—A Global Overview

For indirect vision devices, the regulated fields of view are specified as areas of the ground that need to be covered [8]. In some cases it is also necessary to see the horizon which implies that a vertical field of view is also specified. These specified fields of view correspond to minimum requirements for what should be visible in the most demanding use cases. Think for example about the folded vehicle combination illustrated in Fig. 1 where the driver clearly requires the field of view provided by the Class IV wide-angle mirrors to cope with the situation.

Over the years, the regulated fields of view have gradually increased in size. When traditional mirrors have been relied upon, additional classes of mirrors have been added and a lower glass radius of the existing mirrors has been accepted. Although benefits outweigh, these approaches have also resulted in negative consequences like more extensive obstructions of the direct vision views, from added mirror housings, as well as the condition that indirect vision views can be hard to interpret, due to high distortion when smaller glass radii are needed. For the traditional technology this means there has to be a balance between mirror size that blocks direct vision and the minimum mirror glass radius which can be used to cover the regulated field of view without causing too much distortion.

5.2 Minimum Areas on Ground and Vertical Fields of View

5.2.1 Minimum Fields of View According to UN Regulation No 46 [8]

The UN Regulation No 46 specifies areas on ground in relation to a driver's defined ocular point and the exterior sides of the vehicle as shown in Fig. 20. The specified areas must be covered by the indirect vision devices.

As can be seen in Fig. 20, the main Class II mirrors ask for a more narrow area that extends rearwards all the way to the horizon. In practice this corresponds to an additional rearwards vertical field of view up to the level of the driver's ocular point.

From Fig. 21 it can be noticed that the wide-angle Class IV mirrors ask for a considerably wider area on ground that nevertheless is limited rearwards in length. This limitation in length corresponds to the most common maximum lengths allowed for vehicle combinations.

5.2.2 Minimum Fields of View for Commercial Vehicles in Japan [9]

A comparison is provided by looking into the Japanese vision regulation No 44 which specifies areas on ground in relation to the exterior front, side and rear. In this

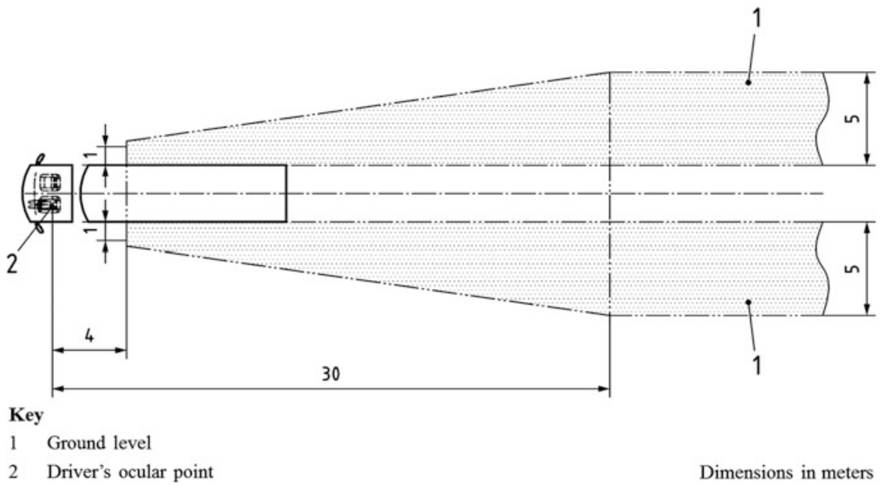


Fig. 20 Area on ground to be covered by Class II mirrors [8]

case the areas can be covered by a combination of mirror types and direct vision through the windows. There are slightly different requirements for light trucks and buses on the one hand and heavy trucks on the other hand. These differences are shown in the following two Figs. 22 and 23.

A lot of focus is put on the front and the passenger side, and the heavy trucks have considerably tougher requirements than the light trucks and buses. In addition to these areas on ground, there are also requirements stating that objects of a certain height must be visible when located at certain positions around the vehicle front and side. However, these requirements regulate the close-up vision rather than the rear-view vision, which is the topic of this book.

6 Certain Driving Situations that Require Different Fields of View

As it was described in the previous section, there are certain use cases which require that larger areas are covered by the indirect vision devices. In the case of traditional mirrors, head movements are normally sufficient to achieve such larger coverage [4, 7]. Figure 19 illustrates this opportunity where the change in the driver's head position actually corresponds to a panning of the view seen via the mirror surface.

In a vehicle which has CMS to provide the indirect vision views, the needs to cover larger areas in certain driving situations are equally important. There are however alternative ways in which they can be provided.

As outlined in Sect. 4, commercial vehicles come in a vast range of configurations. This variation means that it is not possible to decide on one level of

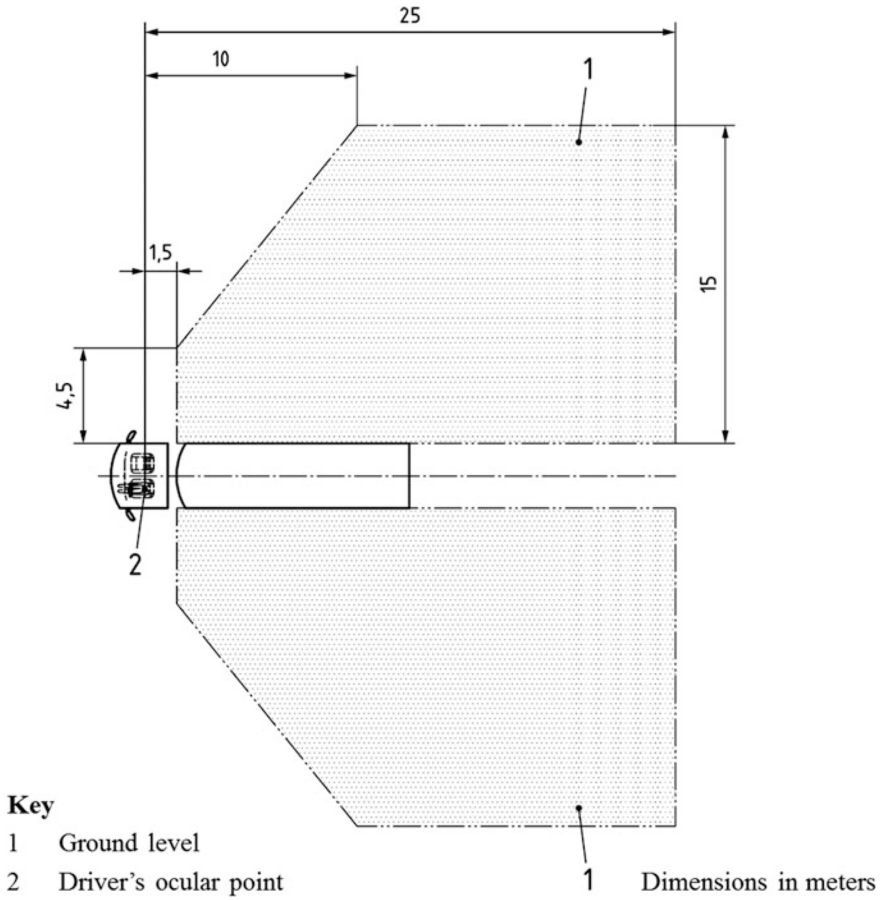
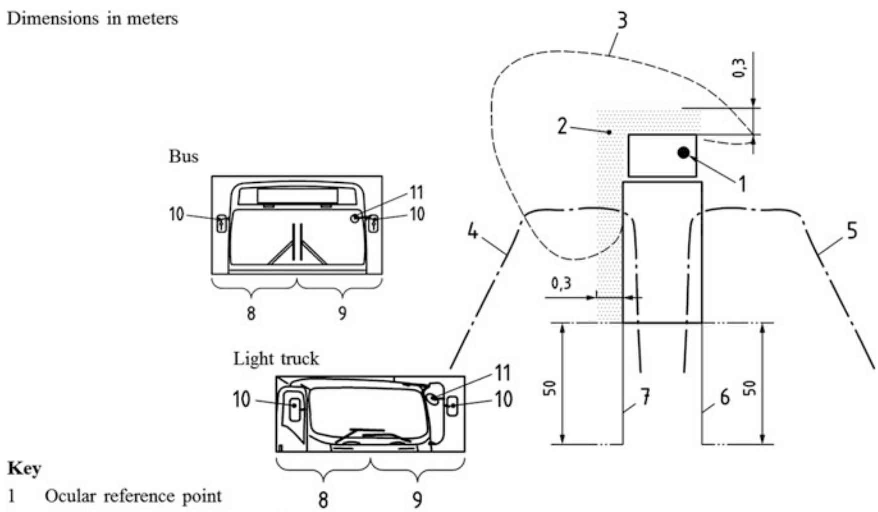


Fig. 21 Area on ground to be covered by Class IV mirrors [8]

requirements that will be appropriate for all configurations. Some attention has to be paid to the actual configuration and the transport task of the specific vehicle [1].

However, as already referred to, it is perfectly viable to divide the available vehicle variations into a limited number of categories where each vehicle category benefits from meeting a certain level of requirements [1]. This approach enables the vehicle manufacturer to come up with the minimum requirements that the CMS has to fulfil. When the needs in one category of vehicles have been focused, the same CMS can then also be used for other vehicle variations that fall inside the same category and that have the same or lower needs for indirect vision views. The purpose is to always secure that the CMS will on all occasions perform at least as well as the traditional mirror system that it replaces.

Based on this approach, it has been possible to establish compulsory instructions in ISO 16505 that outline the procedure how to establish requirements



Key

1 Ocular reference point
2 Field of view for passenger side
3 Front under mirror field of view ($r_{mirror} \geq 0{,}15$ m) for passenger side
4 Passenger side main mirror field of view ($r_{mirror} \geq 0{,}6$ m)
5 Driver side main mirror field of view ($r_{mirror} \geq 0{,}6$ m)
6 Outmost point of driver side
7 Outmost point of passenger side
8 Driver side
9 Passenger side
10 Main mirror
11 Front under mirror

Fig. 22 Area on ground to be covered for light trucks and buses according to the Japanese vision regulation [9]

corresponding to the needs within one vehicle category rather than defining the individual requirements themselves [1]. This procedure should be used to establish the size and behaviour of the provided fields of view. Thus, focus is put on quantifying the fields of view that are needed for the studied vehicle in the use cases that are known as critical. The main critical driving-related use cases are provided in Sect. 3.2.

The requirements stated in this way concern: the establishment of the required size of the changed fields of view; in which way they are allowed to change in terms of changed magnification and temporary transitions; and that the driver is able to fully understand how the system behaves when showing the changed fields of view [1]. When it is secured that the requirements can be adapted to the needs within a certain category of vehicles and at the same time remain open to different solutions for the provisions of changed fields of view, there is maximum flexibility for the vehicle manufacturer to design optimal and cost-efficient solutions while the minimum level of requirements are still being met.

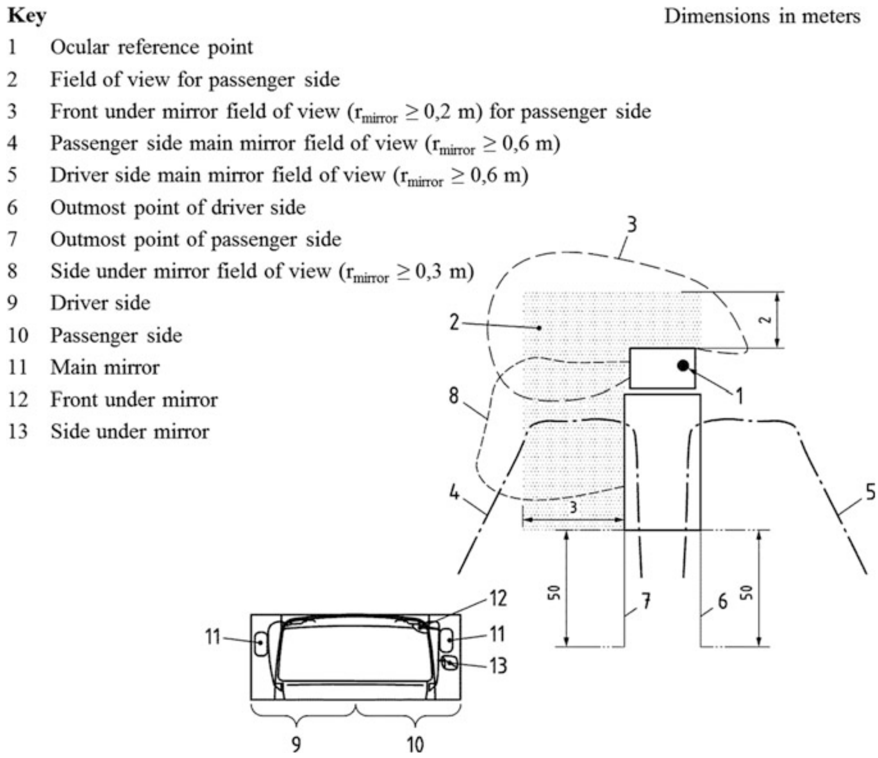


Fig. 23 Area on ground to be covered for heavy trucks, according to the Japanese vision regulation [9]

A brief list of possible solutions that can fulfil the same level of requirements are (1) oversize fields of view that would not need to be changed; (2) temporarily changed fields of view in terms of magnification; and (3) temporarily changed fields of view in terms of panning. If the panning solution is chosen, the panning could either be achieved from automatic recognition of the vehicle conditions corresponding to a use case where a changed field of view is required or from designing the system so it works similar to traditional mirrors in that driver head movements will still result in a changed field of view. But there are of course also a lot of other potential solutions that nevertheless meet up with the minimum level of requirements.

6.1 Procedure for Establishing the Required Changed Fields of View

For special driving situations corresponding to any of the aforementioned critical driving use cases, the purpose is to support that the drivers can keep track of the external boundaries of their commercial vehicle in relation to the surroundings, and that they are also aware of the other road users and their intentions.

As it has already been described, drivers of vehicles with traditional mirrors use head movements to increase the field of view temporarily in order to cover the area that needs to be observed [4, 7].

If the target when designing a CMS is to make sure that it performs at least as well as the traditional mirror system it replaces, it will then become natural to study the level of changed fields of view that the traditional mirror system can provide in a vehicle that typically represents a certain category of vehicles [1].

Such a study has been made in the case of a vehicle consisting of a truck tractor with connected semitrailer [4]. As illustrated in Fig. 18, this vehicle represents a category which has a higher level of requirements for the provided fields of view than other vehicle combinations. In these studies, information was collected from ten experienced truck drivers about both the extent of head movement needed in different driving situations and about which type of mirror that was used when the head was moved. This vehicle was equipped with the main Class II mirrors and wide-angle Class IV mirrors with the fields of view described in Sect. 5.2.1.

In Table 1 below, the main results are provided.

Furthest to the left, the driving situation is provided. The following four columns provide information about the direction in which the head movements were made. The sixth column indicates the average and maximum head movements measured for the participating truck drivers. In the next column the average rate benefit from achieving changed fields of view by moving the head is given according to a five degree rating scale. The rightmost column provides input about the extent to which each type of Class II or Class IV mirrors was used and if their position was on the driver or passenger side.

Table 1 Examples of head movements in special driving situations for a truck tractor and semi-trailer combination [4]

Direction/ situation	Forwards (%)	Rearwards (%)	Left (%)	Right (%)	Average/max (cm)	Benefit (5-degree scale, 1 equals poor and 5 equals excellent)	Mirrors used
Round-about	83	17	50	50	17/30	4	83 % Class II driver side
							67 % Class II passenger side
Reversing	100	0	71	71	33/50	4.8	100 % Class II passenger side
							57 % Class II driver side
Turning	75	0	50	50	17/20	4.5	75 % Class II passenger side
							50 % Class IV passenger side

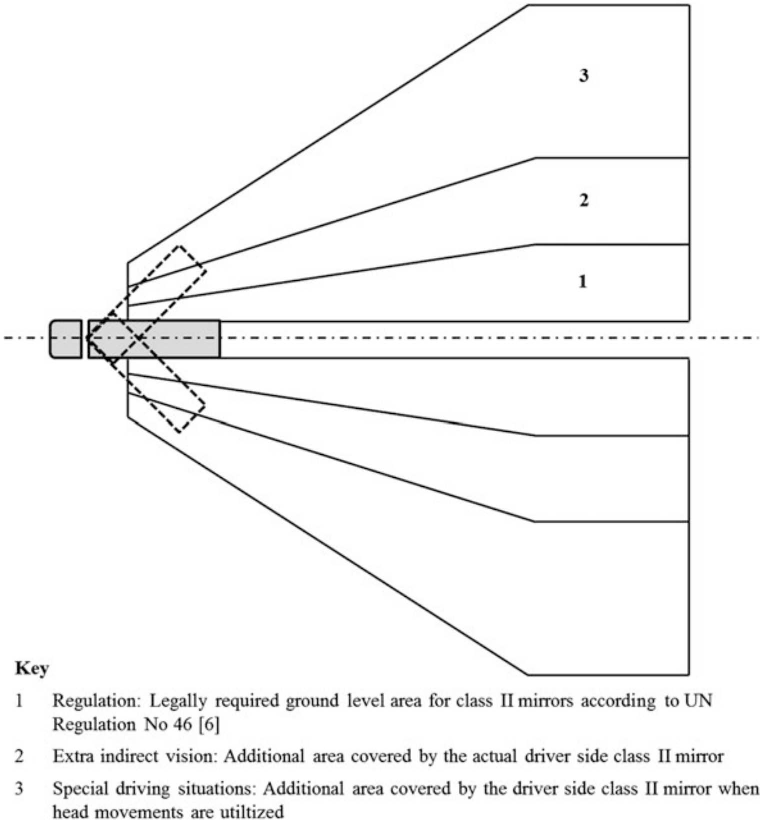


Fig. 24 Relation among the provided fields of view via the main Class II mirrors

The conclusions that could be made from the studies were that Class II mirrors were used more together with head movements except the use case of turning around corners where also the use of the Class IV views was increased from necessity. The preference for using the Class II mirrors can be explained by the fact that they provide views with larger magnification and less distortion which makes it easier to understand what is seen. In addition, the results in the table underline the fact that drivers often have big problems to overview the passenger side when traditional mirror systems are used. Here CMS can provide new opportunities for reducing such drawbacks thanks to image processing or the choice of technical solutions.

From the collected information it was possible to make a drawing to demonstrate the gain in fields of view that the head movements contributed to for the most interesting Class II type of mirrors. In Fig. 24 the total area achieved from additional head movements is shown in relation to the minimum area required by regulation from a fixed eye point as well as the area actually covered by the provided traditional mirror class. The important thing to point out is that the size of the additional area thanks to head movements is quite big compared to the regulated and actually provided areas. This helps to underline the importance of providing changed fields of view tailored to special driving situations.

In this way any vehicle studied can represent a certain vehicle category with need for a specific level of changed fields of view in the critical driving situations that it will encounter.

6.2 Aspects of the Changed Fields of View in Need of Requirements

The previous section brushed at several important aspects related to the changed indirect fields of view that the drivers of commercial vehicles are dependent on. Obvious examples of aspects are the total area covered by the changed field of view as well as the degree of magnification which is required in the changed field of view in order to facilitate the driver's use of the perceived information. This section will briefly describe the total number of aspects that need to be considered according to ISO 16505 [1].

6.2.1 Total Size of the Changed Fields of View

The first thing to consider is the size of the changed field of view so that the size can cope with the worst case vehicle in the most demanding driving situations [1]. As outlined in the example in Sect. 6.1 there are driving situations that require that quite large areas are covered. However, the entire extent is not needed in all ordinary driving conditions. Therefore the CMS should provide such maximum areas, but smart control can make sure that they are displayed only when needed. With traditional mirrors the manufacturers of commercial vehicles normally provide extra fields of view in addition to what is required by regulation. This contributes to a more gradual transfer to changed fields of view by head movements. The CMS can either be manually controlled by the driver or it can automatically utilize vehicle conditions that are typical for more complicated driving situations. In this way, changed magnification, panning, or a combination of the two, can be used to cover the total area that is needed.

6.2.2 Area Coverage by Changed Magnification and Panning

The second aspect to consider is how the changed field of view can be provided. If there were unlimited space available to place monitors inside the vehicle, there would in fact be no need for any changes of field of view. However, the limit for

Fig. 25 Example showing an expanded field of view where the outer part has a different magnification factor [4]

what is reasonable as support for safe driving is reached quite soon [4]. Either the monitor will obstruct large additional areas of direct vision (as illustrated in Fig. 1) or it is necessary to place the monitors in locations that are awkward to overview while driving.

Therefore, the two main remaining means to increase the area of the shown field of view are either to change the magnification so that more can be shown within the same monitor size or to pan the field of view by rotating the camera to make visible more of the adjacent parts next to the original field of view [1].

Figure 25 illustrates an example of what is to be considered in case the necessary magnification is changed only in part of the displayed field of view. It is important that the driver can still interpret the view easily. To realize this, variations in magnification should appear continuously without for example breaking perspective lines [1]. It is also important that the monitor view behaves the same way as traditional mirrors if there is a variation in magnification and related distortion. An example of such preferred behaviour is the phenomenon that the magnification is highest closest to the vehicle side, and lowers as more external areas are viewed [1]. In addition, the minimum and average magnification requirements should also be fulfilled according to ISO 16505 [1] and UN Regulation No 46 [8].

The alternative approach is to use panning. As a matter of fact this can avoid new problems to fulfill magnification requirements. As has already been pointed more than once, the driver actually applies panning already with traditional mirrors since this is the effect when the head is moved. A CMS can, however, provide such in a

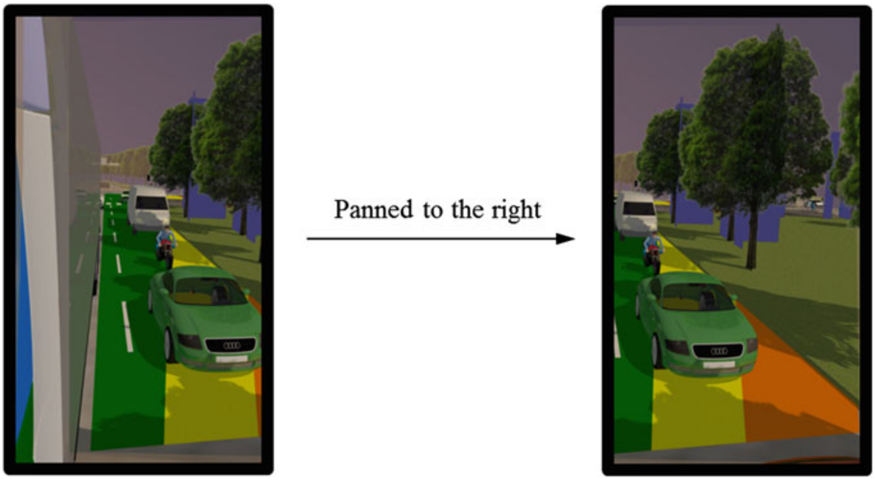


Fig. 26 Illustration of panned fields of view

more optimized way if sufficient information is available about when the changed fields of view are needed. The use of traditional mirrors still relies on the level of experience a driver has of the gains that can be made from the head movements and it requires that the individual driver is able to move freely.

Figure 26 illustrates what it can look like when the main Class II field of view is panned to the right. This corresponds to a changed camera direction or that another part of an oversize view from a fixed camera is shown by means of image processing.

The provision of useful panning will consequently imply changed fields of view at the same speed of transition that a driver can achieve with ordinary head movements in combination with traditional mirrors [1]. If this transition speed is not achieved, there are additional requirements that have to be investigated: either the entire regulated field of view should still be covered also after panning or there is clear evidence that the area without coverage will not be of any use in any of the specific driving situations [1]. As an example, think of the condition when a vehicle is jack-knifing, as illustrated in Fig. 18. Then it can be questioned if the driver is helped by seeing what is underneath the folded trailer part of the vehicle.

6.2.3 Timing Considerations and Changed Fields of View

In the previous Sect. 6.2.1, it was observed that the maximum changed fields of view will be useful only in the specific driving conditions that their definitions correspond to. Therefore it is important that the CMS returns to the originally shown views as soon as the driving condition linked to a certain use case do not apply any longer [1]. See also driving related use cases in Sect. 3.2.

Therefore the identification of time periods after which the changed field of view should return to the original will depend on how far the actual driving situation and use case have proceeded. In real life there will be triggers that depend on vehicle conditions as well as environmental conditions that can be utilized to control the changing fields of view. The triggers related to the vehicle conditions depend on factors such as vehicle speed and how far through a turn the vehicle has proceeded. Triggers depending on the environmental conditions relate to where the surrounding infrastructure and/or road-users are situated in relation to the vehicle itself. As the example in Fig. 18 illustrates, if the rear right corner of a trailer is close to the foundation of a bridge, the driver wants to know the distance until the full vehicle length has passed the concrete foundation. In the same way a driver wants to keep track of a wider field when he or she drives at lower speeds in dense city traffic where many other road-users are passing in and out of the indirect fields of view. On the other hand, when the driver leaves the city area for a rural area with sparse traffic, the wider fields of view are not useful any longer.

With traditional mirrors the needed temporary transitions of such changed fields of view are handled by moving the head to achieve short term changes and by adjusting the mirror orientations to achieve longer term changes. When a CMS is used, the driver can in the same way be provided with a control device which exclusively helps to achieve manual adjustments corresponding to longer term changes [1]. For the shorter term changes, other solutions apply instead. One can be an indirect manual activation that facilitates that automatic changes can take place based on a vehicle condition [1]. An example can be that the panning function of the CMS is not activated until the indicator is used, but the actual panning function will still depend on the subsequent steering manoeuvres in combination with the information about at what speed the vehicle is driven.

This indirect manual activation can support the necessary short term changes in some driving conditions and at the same time help to prevent unwanted changes in other driving conditions. Think for example about the use case when a complex vehicle combination (like any of the ones to the right in Fig. 16) has to be reversed into a narrow spot as the use case in Fig. 12 illustrates. In this case, the driver will most likely need to make several position adjustments by going repeatedly back and forth. If the changed fields of view had only depended on automatic control related to the amount of steering in combination with changes between reverse and forward gears, this would be very annoying.

However, in many other more ordinary driving conditions, fully automatic trigging for both activation and deactivation of the changed fields of view work fine [1]. One important aspect that must be remembered is that the changed field of view will ultimately need to be returned to the original after a certain period of time so that the legally prescribed areas can be secured with all related requirements fulfilled (such as magnification, etc.) [1].

More recommendations about requirements for time periods of activation and deactivation of the changed fields of view are found in ISO 16505 [1].

6.2.4 User Interaction and Interface When Changed Fields of View are Provided

A basic requirement for the interaction with the CMS is that the appearance of changed fields of view does not surprise the driver. When the driver activates the change manually, he or she knows in advance to what extent that manual activation will change the field of view. In this case, the maximum extent of the changed field of view can be shown immediately [1]. The driver will still be prepared.

On the other hand, if the CMS automatically changes the field of view, that change must come gradually so that the driver has time to take in what is happening and keep up a continuous understanding of the scene [1].

As the previous section indicated, there can also be situations where an automatic control will be counterproductive to the use case. In such cases, it is beneficial if the driver can simply turn off the automatic control or lock the changed field of view to a certain degree of change that corresponds to what is needed in the specific use case.

In addition to these basic requirements for the interaction with the system, there can also be other ways to support the driver in the design of the interface. Here the switch into CMS technology from the traditional mirror technology provides new opportunities. To better understand where one's own vehicle is heading in relation to the surroundings and other road users, it is possible to demonstrate the trajectory of which path the vehicle will take if the current motion continues. This is of particular importance in the case of large and more complex commercial vehicles [6].

Another example of how camera views can be enhanced in relation to changed fields of view is if the maximum extent of the changed view is only provided thanks to reduced but uniform magnification. In this case, the monitor will include a lot of detailed information indicated in smaller size than in the original field of view. Under such circumstances it would be beneficial if the known area of interest according to the use case could be highlighted for the driver, particularly important aspects that should not be left unnoticed [5]. The following example illustrates this: the rear end of the driver's own vehicle and its outer boundaries could be rendered brighter in relation to the surroundings at the same time as other strong light in areas of less interest for the particular use case is dimmed down.

With the collected information on what is required in the actual use cases, it would be possible to tailor the monitor interface in the manner described above.

In addition, it could also be investigated how the user interface can be improved by providing added information about the changed field of view. Instances of such information could include what stage of the changed field of view that has been reached or in what way the changed field of view is provided. In this way simplistic scales together with lit symbols could, for example, demonstrate how much of the totally available changed field of view that is currently shown and if what is shown is due to panning, change in magnification, or both.

Apart from the more advanced opportunities that CMS can provide, it is important that the system returns to the original fields of view when the vehicle is parked [1]. Next time the driver starts the vehicle, the ordinary, most useful, recognized and safe fields of view should be shown.

7 Aspects Behind Indirect Vision View Integration in Commercial Vehicles

7.1 Background of Actual Mirror Positions on Commercial Vehicles

The specific usages of alternative types of vehicles in different markets have had as a result that traditional mirrors are provided according to two distinctly different installation concepts [1].

The first mirror installation concept implies that the mirror is mounted to the side of the vehicle; hence the side-mounted mirror. In this case, the mirror is located closer to the driver eye-point and it is viewed through the side window.

The second mirror installation concept instead implies that the mirror is mounted to the front of the vehicle; hence the front-mounted mirror. In this case, the mirror is located further away from the driver eye-point and it is viewed through the windscreen.

These two mirror installation concepts are illustrated in Fig. 27.

The significantly different distances between the mirror glass and the driver eye-point for the two mirror installation concepts correspond to distinctly different levels of magnification [1]. This correspondence has to be considered when a CMS is provided to replace the traditional mirrors on a vehicle. The minimum requirements to be fulfilled are different for the replacement of side-mounted compared with front-mounted mirrors. The detailed minimum requirements are provided in ISO 16505 including Annex A [1]. It is also recommended when a corresponding CMS is provided that the magnification which corresponds to the actual maximum

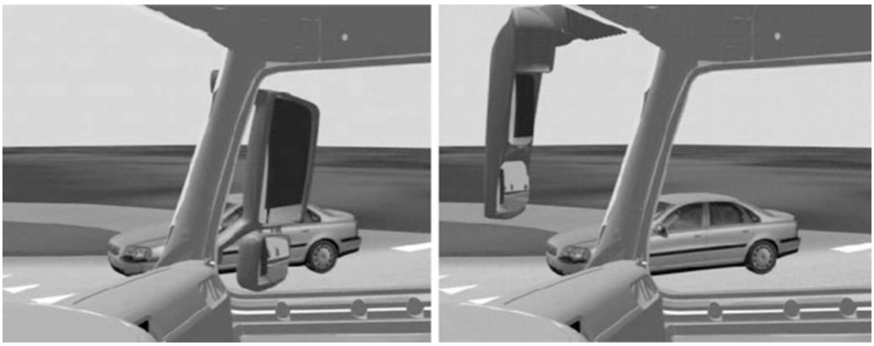


Fig. 27 Difference between side-mounted (*left*) and front-mounted (*right*) mirror concepts

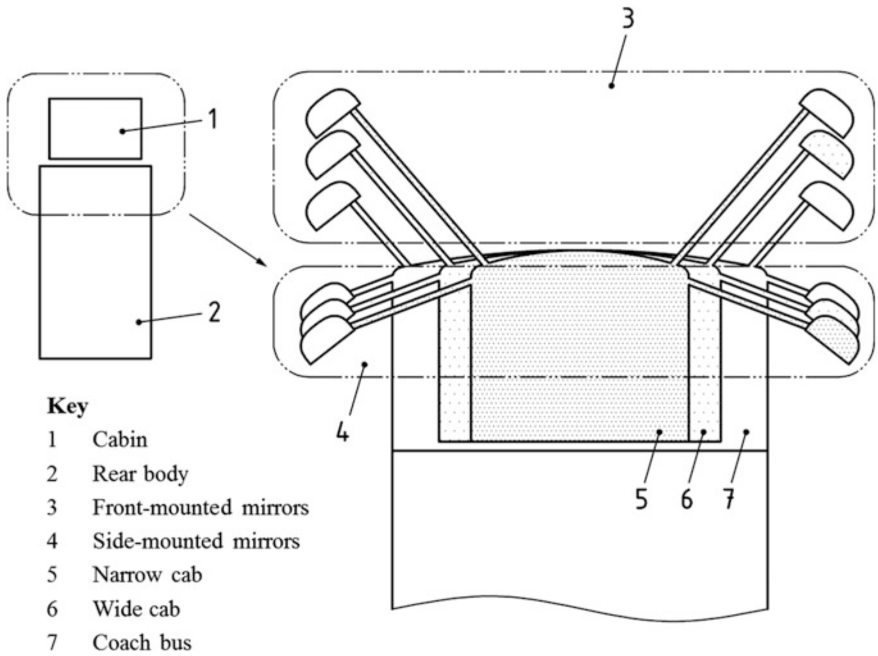


Fig. 28 Relation between mirror concepts and typical commercial cabin widths [1]

distance of the mirror concept to be replaced is used. See also the valid UN Regulation No 46 [8].

Figure 28 further explains the consequences for these main mirror installation concepts including the relations to the width of both the cab and the vehicle in total (which can differ).

As can be understood, the choice of mirror installation concept has a big impact on the viewing conditions for traditional rear-view mirrors. A side-mounted mirror is viewed through the side windows whereas the front-mounted installation is viewed through the windscreen. Besides, the integration of any of the two installation concepts will have a direct impact on how the vehicle can be driven within the existing infrastructure. Depending on the choice of installation concept, the mirrors will protrude outside the vehicle at different locations. Thereby the occupied space that the vehicle needs to be manoeuvred in the traffic will differ. As a consequence the driver will have to adapt his or her driving behaviour accordingly.

When a CMS is applied, the physical connection between the mirror location, to cover the required field of view, and the mirror position, that enables the driver to see the required field of view, is broken. This opens up for some significant improvements of viewing conditions but also of the obstruction of the direct vision by the installation. However, the need to extend camera installation arms so that they can cover the required field of view will still be similar to the corresponding system of traditional mirrors.

The chosen type of mirror installation concept and the relation between cab width and total vehicle width result in different general levels of magnification and resolution within the indirect vision views. The choice of installation concept and the width relations also heavily influence the blind spots caused by the mirror housings.

7.2 *Considerations for Positioning Cameras onto Commercial Vehicles*

Positioning the cameras onto the vehicle exterior should consider five main areas of requirements [4]:

A. Provided fields of view
B. Positions compatible with different vehicle variants
C. Driver's understanding of depth and speed
D. Negative effects from direct light into the cameras
E. Problems of soiling

7.2.1 Provided Fields of View

The fields of view which the chosen camera locations provide must meet with the regulation requirements and driver needs for additional fields of view in the actual driving conditions with the specific vehicle [4, 7]. Both horizontal and vertical fields of view must be supported. See driving related use cases in Sect. 3.2 and minimum areas on ground and vertical fields of view in Sect. 5.2 for additional information.

7.2.2 Positions Compatible with Different Vehicle Variants

To secure solutions that are optimized regarding performance, quality and cost, it is a big benefit if the same camera positions can be used on all variants that the vehicle is offered in [4]. As explained in Sect. 4, commercial vehicles come in a big variety of configurations that are specifically adapted to the type of transport each vehicle is intended for. This variation corresponds both to different sizes of the part of the vehicle that carries the load (both goods and people) and to different sizes of the driver compartment (e.g. opportunities to live in the cab or take additional passengers). Cameras should therefore be placed in exterior areas that are common among different vehicle variants. The chosen camera positions will still have to support the fields of view needed by the vehicle variant that occupies the biggest space on road as well as in relation to the category of load—goods or people.

7.2.3 Driver's Understanding of Depth and Speed

The driver's ability to understand the camera views is to a great extent decided by the chosen camera positions [4, 7]. Different camera heights above ground give different perspectives of the surroundings and they contain information that is easier or harder to grasp (e.g. potential object obstructions and the possibility to determine relations between objects).

Here it should also be remembered that drivers grow accustomed to the views they get from their eye positions when using traditional mirror systems. Placing cameras at other levels above ground will cause unusual views that are hard to interpret.

In addition to this, the camera locations will also govern the necessary level of optical distortion for covering the required fields of view [4]. A low original level of distortion from the camera will simplify the possibility to provide straightforward camera views to the driver without the need for additional image processing.

7.2.4 Negative Effects from Direct Light into the Cameras

Image quality can be reduced if direct light can reach into the cameras [4]. The main surrounding light sources that can cause problems are low sun positions, headlamps of other vehicles, as well as adjacent street lights. Results can be blooming where light sources spread out over a larger view area or limitations to the dynamic range of the camera from which the darkest or brightest parts of the image will be represented as under- or overexposed.

7.2.5 Problems of Soiling

Soiling of the camera lens will directly affect the image quality of the CMS [4]. In addition to this, the way in which the camera is installed onto the vehicle can also cause problems with soiling of windows and thereby reduce direct vision [4]. Soiling is normally categorized as foreign soiling coming from other vehicles and self-soiling coming from the driver's own vehicle.

By positioning cameras further away from the road surface, risks for soiling get reduced. In addition to this, aerodynamic designs of the camera installation and housing can delimit the risk for the soiling of the camera lens as well as the risk that the camera installation causes additional soiling problems. To completely avoid soiling of the camera lens, an active cleaning mechanism can be used.

7.2.6 Common Rule that Solves Most Problems

Luckily most of the requirement areas benefit from the same rule: the higher and further forward that the cameras can be placed onto the vehicle, the smaller the problems of covering the required field of view, optical distortions, surrounding

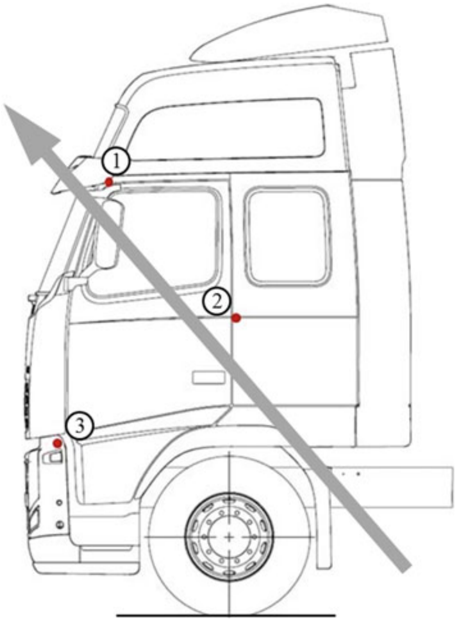


Fig. 29 Benefit of moving the camera higher and more forwards

light sources, and soiling [4]. This common rule is illustrated in Fig. 29 where the numbers exemplify alternative camera locations.

Remaining to consider are the requirements which vehicle variants have and the prerequisite that the camera views should be familiar to the drivers. These requirements will however depend on the designs of the vehicles where traditional mirrors are to be replaced [4, 7].

7.3 Monitor Integration Inside the Vehicle

As explained in Sect. 2, commercial vehicles are specific in the way that a much larger portion of the total vision information comes from indirect vision and in combination with remaining direct vision through windscreen and side windows [1, 2]. This means that it gets of particular importance to support the natural behaviour of drivers as human beings [10]. When integrating CMS monitors within the driver environment, the primary requirements to consider are therefore that the image of the right side field of view should be presented to the right of the driver, and the image of the left side field of view should be presented to the left of the driver [1].

(There could be specific vehicles that are used in special driving conditions at least for part of the transportation work, where another arrangement might be of benefit. In that case, benefits will however need to be demonstrated against drawbacks in general driving use cases including how critical conditions can be handled [1].)

In addition to these basic but very important requirements, there are four more aspects that need to be considered when the CMS monitors are positioned inside the vehicle [4]:

A. Need for eye movements while monitoring vision views
B. Obstruction of direct vision
C. Sensitivity to incoming light
D. Risk of glare and other disturbances in dark conditions

7.3.1 Need for Eye Movements While Monitoring Vision Views

In most vehicles, the main and only reference for positioning visual information is the gazing direction straight ahead onto the road. Also for commercial vehicles this forms an important reference [4]. Depending on use case, it can however be equally important to consider where visual information is placed in relation to gazing directions out of the side windows [4]. In some use cases such directions become essential. Figure 6 illustrates these aspects. Typical examples are reversing at low speed or manoeuvring in dense city traffic.

When both of these positioning references are considered, such monitor positions should be chosen that minimize eye movements both from the most common straightforward line of sight and from direct vision through the side windows in certain driving conditions. Both needs for horizontal and vertical eye movements should be considered.

7.3.2 Obstruction of Direct Vision

The main obstructions of direct vision in vehicles with traditional mirrors are the mirror housings themselves, and the A-pillars on both sides of the windscreen coming second, see Fig. 1. The problem in the case of traditional mirrors is that these obstructions combine and that the two are also dependent on the actual driver position within the vehicle. A shorter driver who sits more forwards gets a different viewing condition than a taller who sits more rearwards.

Transferring to CMS will therefore mean significant general improvements if monitors can be placed so that they do not obstruct additional direct vision in the same way as traditional mirror housings do [4, 7]. Figure 30 shows an illustration of this.

Any additional obstruction that monitors result in other than fixed vehicle structures should be reduced to a minimum. It is an eye-opener to see that A-pillars can cover an entire passenger car in common traffic situations.

7.3.3 Sensitivity to Incoming Light

Surrounding light sources and light levels can cause problems to read the CMS monitor. Light enters through surrounding window areas and can cause problems of

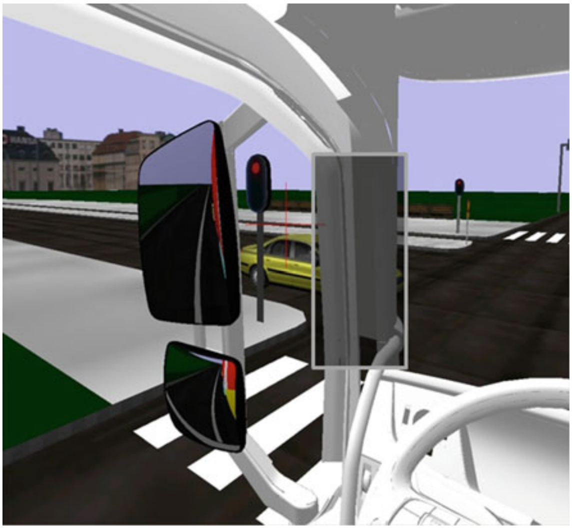


Fig. 30 Example of monitor obstructions in relation to the rear-view mirrors being replaced

glare when it bounces against the monitor surface; of display washout when the light from the monitor cannot compete with the incoming light; of lack of monitor brightness when the surrounding gets very bright; and of disturbing reflections when objects within the vehicle interior shine back into the monitor [4].

Problems can be reduced by preventing light from reaching the monitor surface and by placing monitors in darker areas within the interior. In commercial vehicles the conditions for such solutions can be quite difficult due to needs for more upright driving postures and sufficiently larger side windows to secure appropriate close-up vision. Remaining solutions are to pick monitor angles that minimize the risk for glare and reflections as well as to use anti-reflection treatment of the monitor surface and/or sun-protections around the monitor edges [4].

7.3.4 Risk of Glare and Other Disturbances in Dark Conditions

The brightness of the monitors themselves can also cause disturbing glare in certain driving conditions [4]. The level of disturbance has two reasons: how bright the monitors are in relation to the surroundings and where they are positioned in relation to the common line of sight. The resulting distraction gets aggravated if there are a lot of motions or brightness variations in the monitor views that draw attention away from other visual information within the driving scene.

Several ways exist to counteract such problems of distraction [4, 5]: monitor brightness should adapt to the surrounding light conditions; display technologies and control should be chosen to minimize light leakage and unnecessary variation; monitor positions should be as far away from the straightforward gazing direction as possible; and monitor reflections in side windows should as far as possible be avoided.

8 Key Considerations for Future Development of CMS Replacing Rear-View Mirrors

8.1 Field of View Integration and Simplified HMI

It is a natural consequence to show more than one field of view in the same display, when several mirror classes are to be replaced [1]. Nevertheless, it is still important that separate views with different vision information can be distinguished from one another, and that they are shown simultaneously in a synchronized way. The latter is not only true regarding the basics of providing views, but also when it comes to aspects like colour reproduction and light intensities representing the surrounding conditions.

A CMS is able to handle distortions in a better way than traditional mirrors, both with respect to the difference between different levels of magnification and more similar distortions between driver and passenger side views.

The revolutionary change from today's traditional mirror systems will, however, come when one monitor can provide combined views of more than one field of view, and maybe integrate both rear-view and close-up vision views [11]. When this comes true, minimum requirements according to the ISO 16505 standard will still need to be fulfilled [1]. As views are combined in this way—while still fulfilling requirements for being comprehensible—a lot of redundant information within our current systems of traditional mirrors can be avoided. At the same time the rather complex HMI of today's multitude of mirrors for each mirror class can become much simpler and straightforward [11]. However, it must still be remembered that the driver, the human being, should be able to act instinctively in natural ways, especially in case of critical situations [10].

In parallel with this vision view integration, the opportunities of adding visual warnings in the monitors or adjacent to the ordinary gazing directions will also show the full potential of CMS in comparison with traditional mirror systems. These warnings can also be supported by object detection and more advanced image processing [5]. Also remember what was said in Sect. 6.2.4.

8.2 Opportunity for Common Requirements that Can Be Covered by Either Indirect or Direct Vision

As seen in some markets, there already exist regulated vision requirements that can be fulfilled by indirect vision devices, direct vision or a combination of the two; see Sect. 5.2.2 [9]. There have also been proposals for such combined vision requirements in other markets. For direct vision there exist regulations that specify which areas on ground that should be seen and what sizes of obstructions are allowed [4]. Vehicle manufacturers already have to consider this balance between indirect and direct vision, but common requirements could help to harmonize the approaches.

This chapter has explained that a straight transfer from traditional mirrors to CMS influences both indirect and direct vision regarding which views that are provided and the size of blind spots. In Sect. 8.1 it is also explained that it would be beneficial if rear-view and close-up indirect vision could become more integrated with one another [11].

It would therefore make sense to investigate how common requirements should be arranged in order to meet both indirect and direct vision. Here follow some aspects to consider when establishing such requirements.

The common requirements should be based on typical actual dimensions found within the infrastructure of road systems in concerned markets [2, 5, 7]. In accordance with current indirect vision regulations, the requirements should cover needs for both horizontal and vertical fields of view [8]. Typical dimensions can e.g. be road and lane widths; maximum allowed vehicle lengths, widths and heights; maximum found bridge heights; and standard positions for road signs over the road surface. These infrastructure dimensions will of course vary depending on market, but it would still be possible to focus on the similarities or make requirements with parameter values that can be adapted to the different market conditions.

In addition to this knowledge about actual dimensions of infrastructure within each market, information should be established about the area needs in driving use cases that grasp the combined need for indirect and direct vision. These use cases will then also depend on infrastructure arrangements that can be typical for different markets.

From using combined requirements for indirect and direct vision, vehicle manufacturers could achieve more flexibility when looking into alternative technical solutions and at the same time a higher total system performance level can be secured. The improved total system performance related to vision in commercial vehicles would then also result in benefits for society regarding safety and cost.

8.3 Communication with and Information to Other Road Users

The traditional mirrors have—at least for classes with larger radius—the benefit of providing opportunities for visual dialogue. In practice this means that road users have the opportunity to see each other's faces as a help to understand the mutual intentions. Especially in dense city environments this can constitute a big benefit. With a CMS, this opportunity does not come as an automatic consequence, that is, the drivers of the vehicles cannot have eye contact via a mirror to get the confirmation that both have seen each other.

To secure good communication between vehicles about mutual intentions is however of general interest, and could be explored further as new technical solutions make it possible to exchange information between vehicles. Here new technologies can also be less dependent on traffic situations or weather conditions to secure successful communication of higher quality.

8.4 Available Locations for Camera Installations

It would help if cameras could be placed where they provide the best views of the vehicle and the surroundings. Historically, the vehicle manufacturer has been responsible for the installations on the vehicle, the trailer supplier for the installations on the trailer and the provider of infrastructure for the installations within the infrastructure. If the responsibilities were not split and if the interaction between these parties could evolve, it would be interesting to have cameras mounted also on optimized locations on trailers or at suitable infrastructure locations [6, 7]. This could lead to better views and less need for changed views which might imply more robust systems. The same is valid for sensors of different kinds which could provide information of higher quality if their positioning were not limited to the vehicle only [6]. Figure 31 shows an example of this where the view from a camera at a goods terminal could be transferred to the monitor in the vehicle.

Fig. 31 Camera view from infrastructure that could be transferred into the vehicle [6]

8.5 Developed Content of Driver Tasks

With the increased focus on automation, we can foresee a transition of driving tasks going from being the driver with continuous attention to driving, to becoming an operator who observes the vehicle system and acts when necessary. The same transfer has already taken place within the shipping industry; where as an example mates rely on automated manoeuvring and utilize a multitude of camera views instead of keeping track of the ship position from viewing bridges to overview manual manoeuvres by direct vision. In this context, the introduction of CMS for vehicles and infrastructure could support a change into alternative future developments.

9 Conclusion

It has been established that commercial vehicles have specific needs that need to be considered when Camera Monitor Systems replace rear-view mirrors. Unique use cases have been identified that correspond to additional requirements which will have to be fulfilled to make this new technology work as intended. At the same time the shift of indirect vision technology can provide big opportunities, to the point of actually revolutionizing the vision conditions for the drivers. This includes improvements of the indirect vision views as well as the combination of indirect and direct vision when the limitations related to the physics of traditional mirrors have been removed.

A first target when harvesting the potential benefits of this new technology will be to make significant improvements of the work environment for drivers of commercial vehicles. Correctly applied, Camera Monitor Systems can make it easier to overview these big and often complex vehicles and also the surroundings, including other road-users in critical traffic conditions. These benefits correspond to huge gains both in transport efficiency and in traffic safety: driving commercial vehicles would become less stressful; the interrelation between commercial vehicles and other road users would greatly improve; and it would become possible to perform manoeuvres, within existing infrastructure, which are currently not feasible with traditional mirror technology.

Once the technology shift has taken place, several additional functionalities are foreseeable. They range from new approaches to provide drivers with adapted vision information to solutions for how available vision information can be transferred between vehicles and infrastructure to extend usability in new ways. The more intelligent transport systems supported by these leaps in development result in an increased efficiency of transports and safety for the drivers of commercial vehicles as well as other road-users. This would have clear positive effects not only for transport and traffic but for the entire society.

References

1. International Organisation for Standardisation. ISO FDIS 16505 – Road vehicles – Ergonomic and performance aspects of Camera Monitor Systems – Requirements and test procedures. ISO/TC 22 Road vehicles, Subcommittee SC 17 Visibility, ISO, Geneva, Switzerland (2014)
2. Fitch, G.M., Blanco, M., Camden, M.C., Olson, R.L., McClafferty, J., Morgan, J.F., Wharton, A.E., Howard, H.E., Trimble, T., Hanowski, R.J.: Field Demonstration of Heavy Vehicle Camera/Video Imaging Systems. Report no DOT HS 811 475, Center for Truck and Bus Safety, Virginia Tech Transportation Institute, National Highway Traffic Safety Administration, Washington, United States (2011)
3. Danielsson, H.: Critical Vision and Adapted Driving Postures. Volvo Group Trucks Technology, Gothenburg (2011)
4. Fornell Fagerström, K., Gårdlund, A.: Mirror replacement in trucks. Department of Product and Production Development, Chalmers University of Technology, Gothenburg, Sweden (2012)
5. Wierwille, W.W., Schaudt, W.A., Blanco, M., Alden, A.S., Hanowski, R.J.: Enhanced Camera/Video Imaging Systems (E-C/VISs) for Heavy Vehicles. Report no DOT HS 811 483, Virginia Tech Transportation Institute, National Highway Traffic Safety Administration, Washington, United States (2011)
6. Borre, K., Larsson, R.: Low speed maneuvering aids for long vehicle combinations. Department of Product and Production Development, Division of Design and Human Factors, Chalmers University of Technology, Gothenburg, Sweden (2012)
7. Wierwille, W.W., Schaudt, W.A., Spaulding, J.M., Gupta, S.K., Fitch, G.M., Wiegand, D.M., Hanowski, R.J.: Development of a performance specification for camera/video imaging systems on heavy vehicles. Report no DOT HS 810 960, Virginia Tech Transportation Institute, National Highway Traffic Safety Administration, Washington, United States (2008)
8. United Nations Economic Commission for Europe. Indirect vision devices – Uniform provisions concerning the approval of devices for indirect vision and of motor vehicles with regard to the installation of these devices. ECE R46-04, InterRegs Ltd, United Nations (2014)
9. Japan Ministry of Transport. Rear-view mirrors. 11-1-44, Japan Automobile Standards Internationalization Center, Japan (2013)
10. Danielsson, H., Höcke, E.: Future Layout of Visual Information in Trucks. Department of Product and Production Development, Chalmers University of Technology, Gothenburg, Sweden (2013)
11. Zaindl, A., Zimmermann, A., Dörner, K., Kohrs, C.: Camera-Monitor System as Mirror Replacement in Commercial Vehicles. Automobil Technische Zeitschrift, Germany (2015)

Part II
Fundamentals of Automotive Technology for CMS

Image Sensors for Camera Monitor Systems

Michael Brading, Brian Keelan and Hieu Tran

Abstract Image sensors are a critical part of camera monitor systems (CMS). Photons passing through the imaging lens impinge upon microlenses that focus the light through color filters and onto the sensor pixel photodiodes. Two principal sensor architectures in use today are front-side and back-side illumination (FSI and BSI), the latter of which is just becoming available in the automotive space. Pixel signal integration and readout can be implemented by electronic rolling shutter or global shutter mechanisms; the former dominates viewing applications but the latter can be advantageous for machine vision. The signals are read out by analog circuitry, and converted to the digital domain. The digital signal is processed to produce an image suitable for display or analysis. This process includes steps that: correct for black level and white balance; fill in missing color signals via a demosaic algorithm; spatially process the image to distinguish signal and noise; map the signals to a standard color space; and apply tone-scaling. Critical sensor performance measures for CMS applications include signal to noise ratio (SNR) and modulation transfer function (MTF); the advanced Fourier metric noise equivalent quanta combines these two measures and is useful in understanding task performance. Scene dynamic range in automotive applications can be very high, beyond the capabilities of a single pixel exposure, and therefore special techniques are required. The mitigation of flicker from time-varying light sources, such as fluorescent lamps and light emitting diodes (LEDs), is a current challenge, particularly in high dynamic range (HDR) imaging. To exemplify the practical implementation of image sensors in CMS, a sample system design is presented. Image sensors with integrated safety features support CMS implementations that require ISO 26262 automotive safety integrity level (ASIL) ratings. CMS is an exciting new application, with considerable opportunity for innovation, particularly in the image sensor and processing components. Advances in these areas will contribute to the greater adoption of CMS systems in vehicles in coming years.

M. Brading (✉) · B. Keelan · H. Tran
ON Semiconductor, San Jose, CA, USA
e-mail: michael.brading@onsemi.com

A. Terzis (ed.), *Handbook of Camera Monitor Systems*,
Augmented Vision and Reality 5, DOI 10.1007/978-3-319-29611-1_5

Keywords Image sensor · Camera monitor system · Microlens · Color filter array · Pixel · Back-side illumination · Readout circuitry · Analog to digital converter · Signal to noise ratio · Noise equivalent quanta · Auto functions · Image processing · High dynamic range · LED flicker · Motion artefacts

1 Introduction

The image sensor is a critical part of a camera monitor system (CMS). Charge coupled device (CCD) image sensors have been used in vehicular camera systems since the 1990s. Over the last decade, complementary metal oxide semiconductor (CMOS) image sensors have become widely used in viewing systems, such as backup and surround-view cameras, and in machine vision applications, such as lane departure warning and headlamp control. CMOS sensors, with their on-board processing and lower cost, have come to dominate the market for vehicular vision systems, so we focus our attention on CMOS imagers in this review.

2 Image Capture

We first review the image sensor capture system from light collection by the sensor microlenses to the resultant digitized signal from the analog to digital converter.

We can see in Fig. 1 that light from the scene is focused via the camera lens onto the image sensor. Typically in automotive applications the camera lens is made up

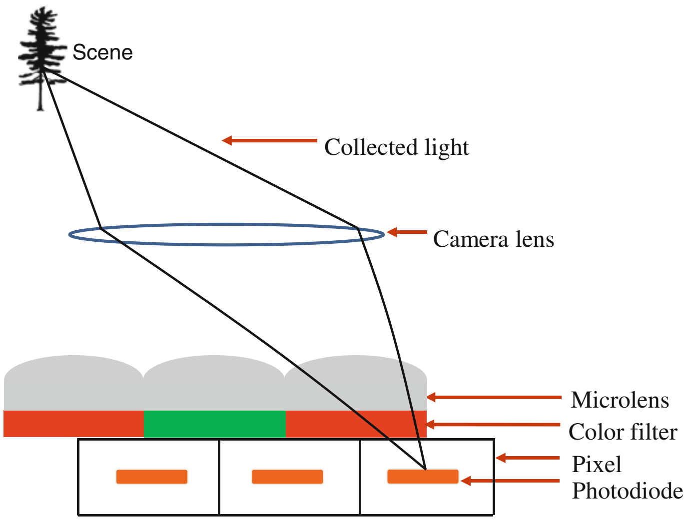


Fig. 1 Optical light collection

of five glass elements. The light is further directed by the sensor microlenses before passing through a color filter, and finally entering the pixel. Upon reaching the silicon photodiode, the photon energy can create an electron-hole pair. Charge created by this action is stored and converted to an electrical voltage and subsequently to a digital signal.

In an ideal system the light emerging from the camera lens would enter each microlens and pixel in a perpendicular fashion, as occurs with a telecentric camera lens. In practical implementations the chief ray angle (CRA) of the light is usually not perpendicular except at the center of the field of view, especially for lenses with a wide field of view. Therefore the microlenses are generally shifted as shown in Fig. 1 to correctly direct the light through the color filter and into the photodiode. The degree of shift varies with the angle from the optical axis (field height) and is chosen to match the chief ray angle profile of a representative lens.

Today image sensors for automotive applications are constructed almost exclusively using complementary metal oxide semiconductor (CMOS) technology. The electrical and physical architectures of these pixels vary widely depending upon the application requirements and manufacturer. Here we will review two of the fundamental construction techniques; front-side illumination (FSI) and back-side illumination (BSI). The FSI construction (Fig. 2) requires that light enters via the microlens and color filter and then travels between the metal layers of the device interconnect before reaching the photodiode. This approach is quite mature and has been highly tuned to optimize performance. As pixel geometries shrink and lens chief ray angles increase with increasing fields of view, it will become advantageous in automotive applications to adopt BSI technology, which has been very successful in consumer electronics products such as smartphones. BSI construction (Fig. 3) moves the photodiode such that it is immediately below the microlens and color filter, and above the otherwise potentially obscuring metal layers. This permits greater angles of incident light without artifacts (such as crosstalk) and higher levels of efficiency in small pixel geometries where the metal interconnect could have significant impact upon the incident light.

In manufacture a BSI sensor is initially constructed in a similar fashion to an FSI sensor. Once the silicon and metallization layers are fabricated a carrier wafer is attached to the front side of the wafer, then the complete assembly is inverted and the back side of the wafer polished until the original wafer is less than 10 μm thick. This allows the light to pass through the back of the wafer directly onto the photodiode. After the polishing step the color filters and microlenses are constructed on the back of the wafer over the photodiodes.

A typical image sensor used in automotive applications (Fig. 4) is constructed from an array of 4-transistor pixels. The pixels are addressed in a row-wise sequential fashion; this method of readout is known as an electronic rolling shutter (ERS). The timing generator and row decoder produce two essential pointers that roll down the image sensor during a single frame time: the reset pointer, which resets each row of pixels; and the sample pointer, which enables the readout of the stored charge. The time elapsed between the readout and the sample is the exposure or integration time of the image. The time elapsed whilst the pointers cycle through

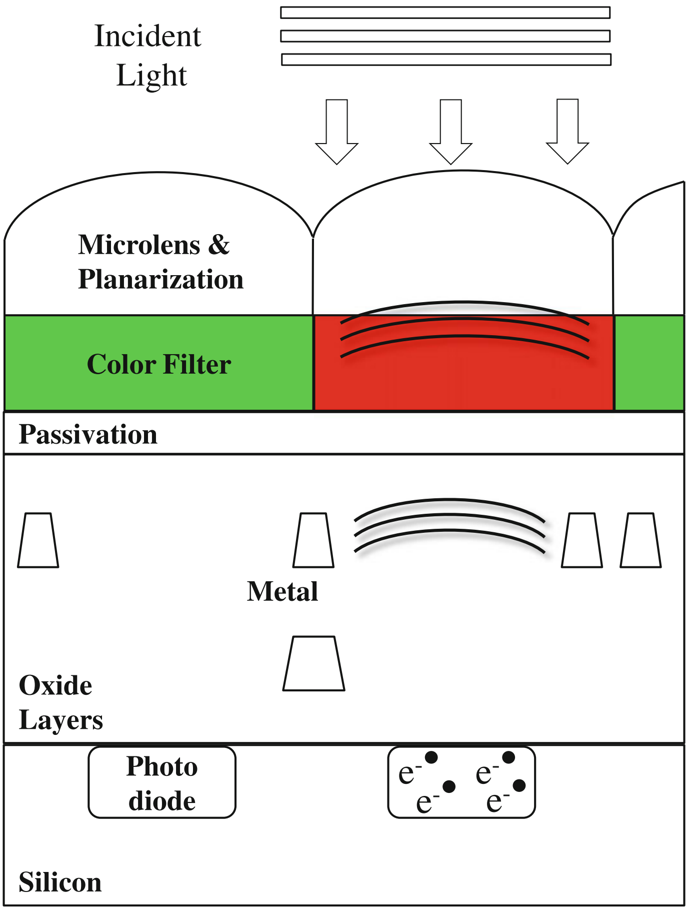


Fig. 2 Front-side illumination sensor

a complete frame readout is known as a frame time. A CMS is likely to operate at 60 frames per second or higher to ensure smooth motion.

In some applications that involve camera shake or high speed motion it is desirable to use a global shutter (GS). This can be implemented with 5 or more transistors. Below we show an example with 5 transistors, which allows the entire array of pixels to be reset at the same instant and then integrated for the same time. The integrated charge within each pixel is then stored and sequentially read out in a similar manner to the ERS system. The advantage of the GS design is that if the camera or the scene being captured are subject to high speed motion, because all points in the image are captured at a single moment in time, there are no motion inconsistencies. In the ERS design the pixel rows are sampled sequentially so that if an object in the scene spans multiple rows and moves during the capture, then motion distortion can be seen. A GS design can exhibit higher dark current than an

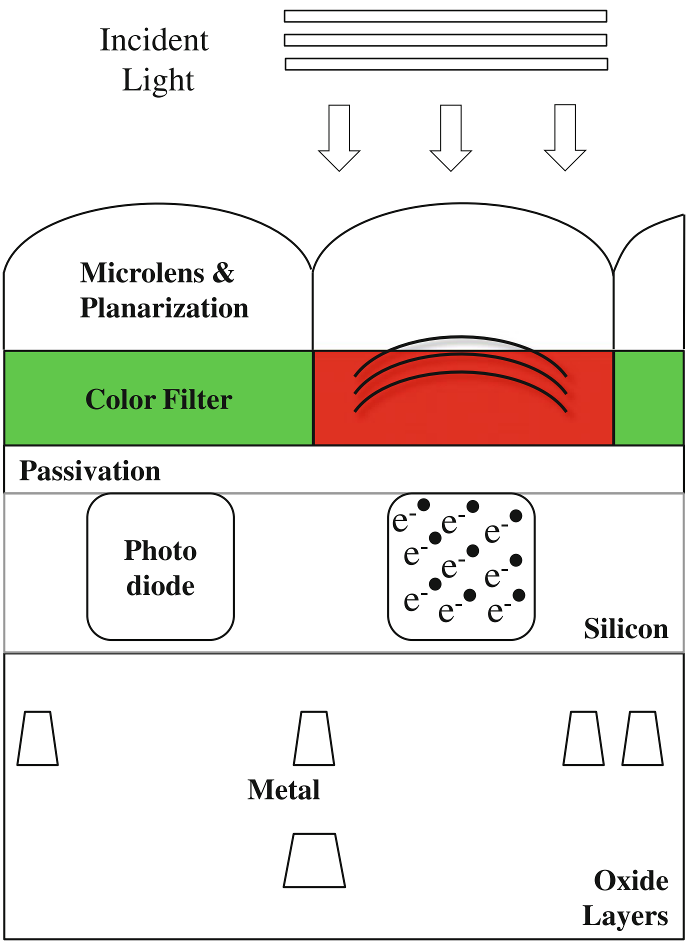


Fig. 3 Back-side illumination sensor

ERS design, and there is the potential for light leakage if the in-pixel storage node is not well shielded from incident light. This contamination of the signal causes a bias that increases with the time between sampling and readout, and so produces a vertical shading effect in the image.

Figure 5a shows an active pixel in a typical 4-transistor (4T) configuration. A 4T pixel consists of a pinned photodiode, a reset transistor, a transfer transistor, an amplifier transistor and an output row select transistor [1]. The 4 transistor architecture provides an excellent low noise pixel and is widely used in automotive image sensors with a rolling shutter operation. In contrast to a 3T architecture, the 4T architecture permits the use of correlated double sampling to eliminate noise caused by variations in the reset voltage of the floating diffusion across the array. The reset noise is a result of the thermal noise of the floating diffusion (FD) and is a

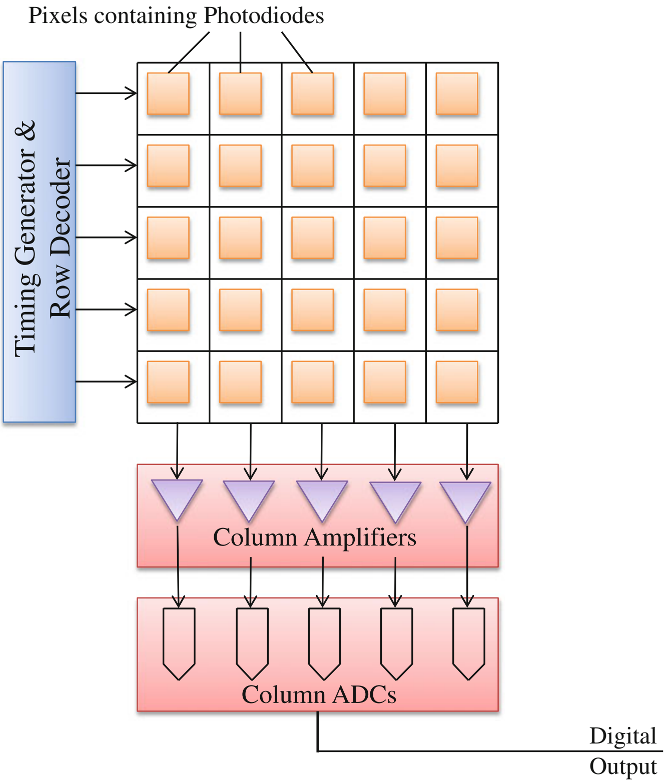


Fig. 4 CMOS image sensor structure

function of the capacitance C of the junction and the temperature T, and so is also known as kTC noise, where k is Boltzmann's constant. Correlated double sampling, as its name suggests, samples both the reset and signal levels during the same readout cycle and provides the difference between them, canceling the kTC noise.

The fundamental 4T architecture is possible because of the use of the pinned photodiode, which ensures that complete charge transfer is made from the photodiode to the source follower (SF) when the transfer (TX) gate is enabled. Without the use of a pinned photodiode the charge transfer could be incomplete and the pixel would exhibit image lag due to the charge retained in the photodiode.

Operation of a 4T pixel is as follows. Recall that two primary pointers are operating within the array, the reset pointer that initiates integration and the sample pointer that ends the integration and samples the accumulated charge. The reset pointer enables both the reset transistor (RST) and the transfer transistor (TX), which charges the photodiode (PD) to the analog voltage (VAA). Electrons generated by incident light then discharge the PD node. After the desired integration time the sample pointer then operates the pixel as per the timing shown in Fig. 5b. With the row select (RS) transistor enabled, the FD is reset and read out, including

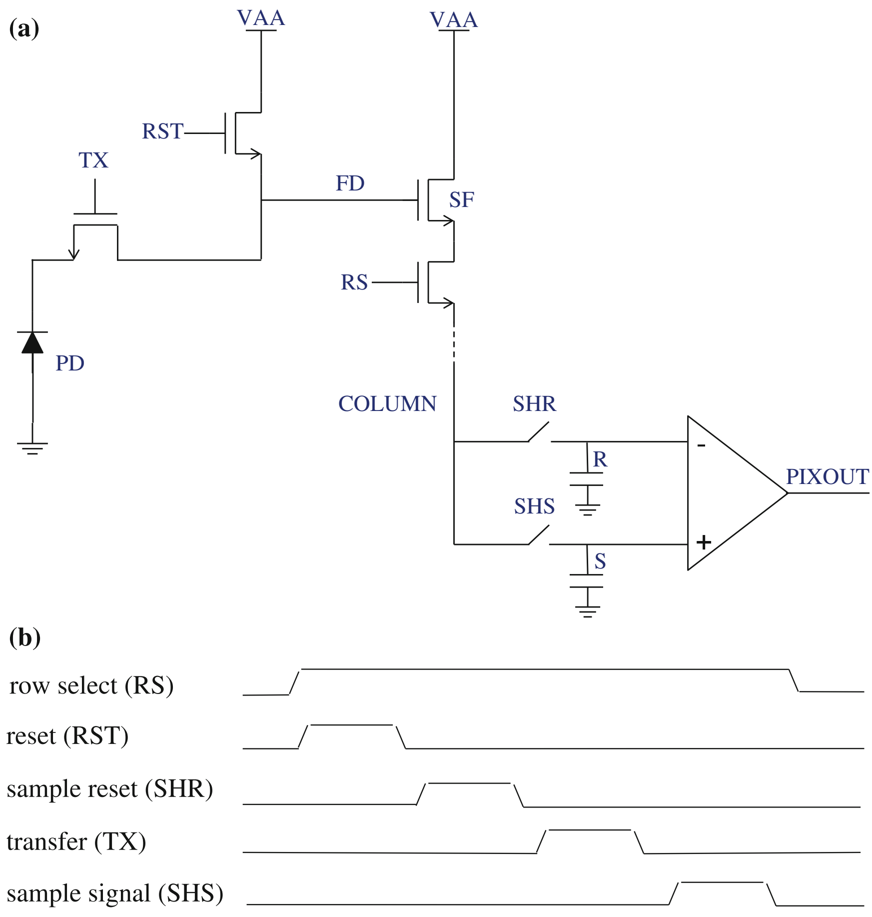


Fig. 5 **a** 4T pixel. **b** 4T pixel timing

any offsets, and is sampled onto the capacitor R. Then the TX gate is turned on and the charge from the PD is transferred to the FD and the FD is read out and stored on the capacitor S. The integrated charge as a result of incident light and dark current is equal to the difference between the signal sample and the reset sample.

There are several performance parameters of a 4T pixel to consider when developing an image sensor array for automotive applications. Here we review three of the most important: sensitivity, dark current and dynamic range. Sensitivity is a measure of the signal generated for a given amount of incident light energy. The signal generated can be measured at various locations in the readout signal chain and thus the signal can be expressed in a number of different units. If measured at the floating diffusion node, the units are electrons; at the column amplifier output, volts; and after the analog to digital converter (ADC), digital counts (also called least significant bits,or LSBs). The incident illumination for visible light sensors is typically measured in lux-seconds (lux-s) and so the sensitivity is commonly expressed as e^{-}/lux-s or V/lux-s.

Dark current is a measure of the charge generated within the pixel during the integration as a result of thermally generated electrons. Dark current typically doubles about every 8 °C and is a physical property of silicon. Careful processing and design of the junction can reduce the dark current but not the rate at which it increases with temperature.

The dynamic range of a pixel is the ratio of the linear full well capacity of the FD node to the minimum detectable signal, usually taken to be the noise floor (the sum of all signal-independent pixel noise sources including dark current). The system dynamic range can be extended beyond that of a single pixel exposure through either temporal (multi-exposure) or spatial (multi-pixel) means, as described later.

A global shutter (GS) pixel is used when there is a requirement to capture the entire image at one moment in time, as opposed to the electronic rolling shutter (ERS) where the image is captured sequentially on a row-wise basis. There are many variants of GS pixel design and here we describe one of the most mature and simple architectures utilizing just 5 transistors. Other variant architectures that utilize 6 or more transistors also exist [2]. In the 5T design shown in Fig. 6 note the additional transistor labeled TX1. The operating sequence is as follows. First the entire array is reset globally, and the image acquisition starts. The readout of the photodiode is also made globally by transferring the charge stored in the PD onto the FD node. The FD node is used to store the charge until such time as it is read out in a row wise fashion in the same way as in a 4T pixel. A key design requirement of a GS pixel is to shield the FD node from incident light to avoid corruption whilst the readout occurs. If the FD node is not adequately shielded one

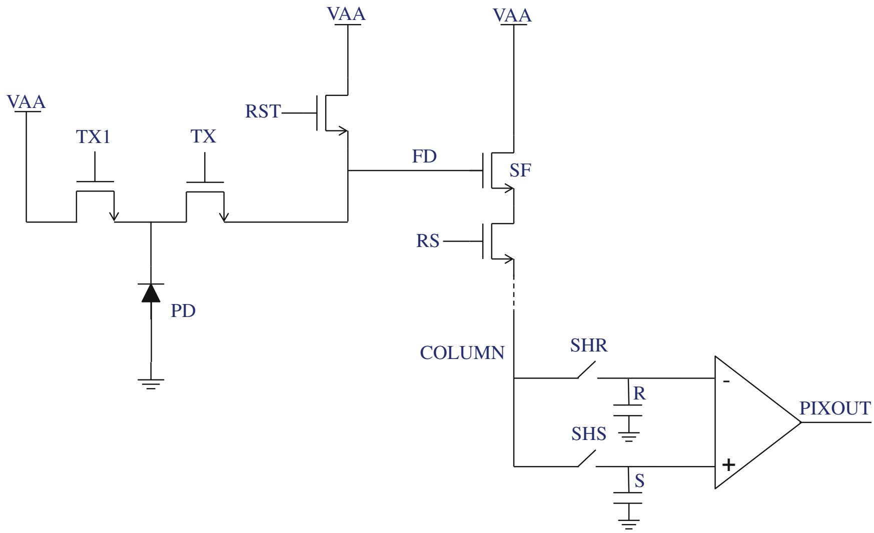
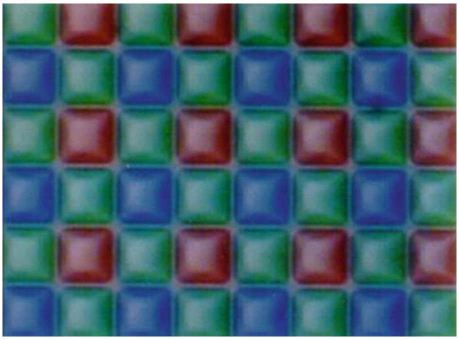


Fig. 6 5T pixel

Fig. 7 Bayer color filter array



can observe vertical shading in the image with the bottom of the image being brighter due to light leakage post-integration and prior to readout. The fifth transistor TX1 is used to ensure that charge accumulating on the PD during readout is drained to the analog supply voltage VAA and does not corrupt the FD node.

A CMOS photodiode is sensitive to all visible wavelengths of light. To construct an array of pixels with channels sensitive to individual primary colors (e.g., red, green, and blue), a filter is placed above each pixel, creating a color filter array (CFA) mosaic. Photons outside of each filter's passband are absorbed, and by adjusting the thickness of the filters, the selectivity can be modified. There are a variety of different CFA patterns, the most popular being the Bayer pattern, which is shown in the photomicrograph of Fig. 7.

Sensor response depends not only on the CFA spectra, but also upon the spectral response of the silicon photodiode. Sensor response is usually characterized by its quantum efficiency (QE), which has units of electrons per incident photon, and so has a maximum of 100 %. Sensor QE is not simply the cascade (wavelength by wavelength multiplication) of the photodiode QE and the color filter transmittance spectra, because the imaging light forms a cone, with different rays traveling different distances through the silicon.

In addition to sensor QE, overall camera response of the different color channels depend upon several additional factors: (1) the scene illuminant spectral power density; (2) the camera lens transmittance; and (3) the infrared cutoff filter transmittance. The combined response from contributions (1)–(3) and the sensor QE is shown for a representative camera in Fig. 8. The assumed illuminant is CIE D65, with the spectral density converted from watts (per unit area per unit wavelength) to photons, so that the units cancel properly with those of the QE. The area under the green curve is highest, implying greatest sensitivity in this channel in daylight. Automatic white balance, discussed in the next section, adjusts gains to compensate for the lower areas under the red and blue curves, to yield an approximately neutral reproduction of the scene.

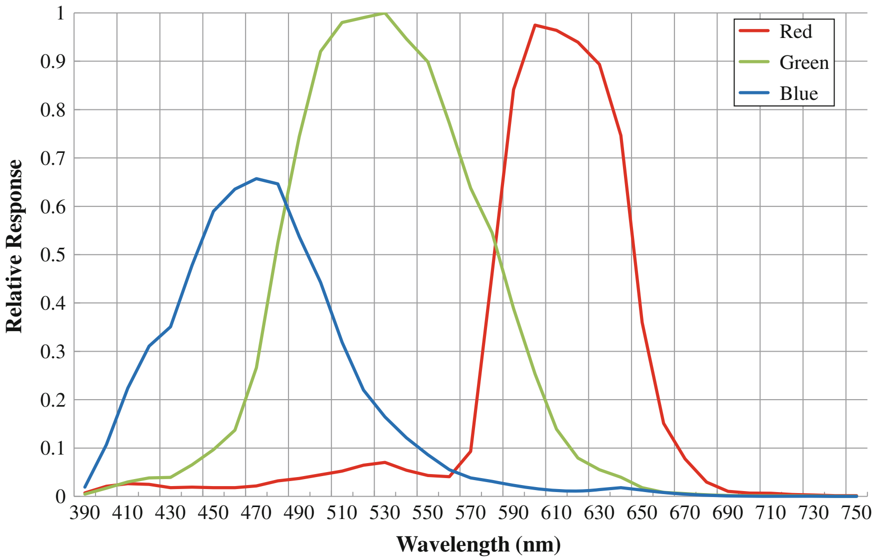


Fig. 8 Camera spectral response functions

3 Image Processing

Figure 9 shows some of the operations in a typical image processing pipeline. There are some variations possible in the ordering of the operations, and a few can be done in either the analog or digital domains, or both (e.g., black level subtraction and white balance).

Black level subtraction zeros the signal in the absence of light, so that the final signal is proportional to exposure. It corrects for dark current, and so the magnitude of the correction varies with integration time (proportional) and temperature (exponential). The correction can also vary spatially for several reasons such as temperature gradients. Sometimes an adaptive (image-dependent and location-dependent) operation is done later that partially compensates for flare, arising principally from lens,

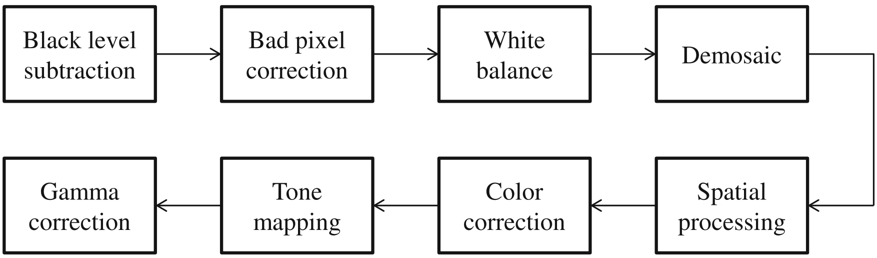


Fig. 9 Image processing pipeline

filter, and sensor surfaces. Flare produces spurious signal in dark areas, like dark current. The flare compensation is essentially a tone scale correction operation.

Bad pixel correction generates estimates of the correct signal values for pixels that are faulty, based upon the values of nearby pixels. Single-pixel defects can typically be corrected satisfactorily on the fly by a digital algorithm. The algorithm ideally takes into account a number of nearby pixels (e.g., a 5 × 5 neighborhood) and maps out edges in the region to avoid interpolating across an edge (leading to blur and artifacts). Cluster defects, in which multiple nearby pixels are faulty, usually require that the location be mapped and stored in memory (such as one-time programmable memory) for the reference of the correction algorithm. Faulty pixels usually are either dark pixels that have little response, or hot pixels that have defects in the silicon structure producing spurious signal, and causing the pixels to appear too bright if uncorrected.

White balance is an operation in which two of the color channels are gained up to match in integrated response to a third channel, so that the resulting image rendition is neutral (meaning that neutral elements of the scene appear neutral in the final image). The necessity for such an operation is evident from Fig. 8, where area under the green channel response exceeds that under the red, which in turn exceeds that under the blue. Without white balance, the image would appear to have an overall yellowish green cast. The necessary white balance gains are a function of the scene illuminant, which often is unknown and so must be inferred. White balance algorithms involve statistical estimation and can be extremely sophisticated; state-of-the-art algorithms have reached a remarkable level of performance in recent years.

An array of filters having different colors, as in Fig. 7, is referred to as a mosaic pattern, and so the process of filling in the missing color channel data at each pixel position (e.g., red and blue values at a green pixel position) is called a demosaic operation. Demosaic algorithms, particularly for the Bayer pattern of Fig. 7 (with a 2 × 2 unit cell of RG/GB), have been refined to the point where there is almost no resolving power loss associated with the mosaic pattern. In a Bayer pattern, the checkerboard arrangement of green pixels, occupying half the area and being present in every row and column, allows a high fidelity capture in the green channel. This is advantageous because the luminance response function of the human visual system, which substantially determines the perception of sharpness and noise, peaks in the green, which is therefore often referred to as the luma channel. The interspersed red and blue pixel values can be used to improve the reconstruction (filling in of missing values) of the green plane. Once the green plane is complete (every pixel has a green value), it can be advantageous to subtract the green values from the sparse red and blue values to make two difference (chroma) channels. The edges and texture of the scene substantially cancel when the difference is taken, so the chroma channels have relatively little high-frequency detail. As a result, they can be interpolated with fewer artefacts—after which the green channel can be added back to reconstitute the red and blue channels.

Spatial processing refers to operations that use multiple pixel values from a local region to revise values of a given pixel. For example, each green pixel value could be replaced by the average of green pixel values over a 3 × 3 region centered on that pixel, to reduce noise. This operation is equivalent to a low-pass filter that will also blur (smooth) any high-frequency detail present in the region. Alternatively, a weighted sum of local pixels involving some negative coefficients can be used to enhance edges, sharpening an image; however, any noise present is amplified to the same degree. Consequently, the key to successful spatial imaging is to use adaptive processing, in which each local region is classified as being dominated by noise (i.e., a flat region, with little detail) or being dominated by signal (edges and texture). Then regions dominated by noise can be smoothed, whereas regions dominated by signal can be sharpened. The relative magnitude of signal and noise depends not only on location within an image but also on spatial frequency. Average scenes have less high-frequency than low-frequency signal content, and the capture system further attenuates high-frequency content because of its modulation transfer function. Consequently, the spatial processing should have a frequency spectrum complimenting the signal to noise profile of the imagery.

To prepare an image for display or other interpretation, several steps are typically taken. First, the red, green, and blue camera responses (RGB) are mapped to a standard color space, the most common example being "sRGB" [3]. Displays with different primaries can be calibrated to provide accurate color under the assumption that the capture response closely approximates the sRGB standard. To transform device RGB to sRGB, a 3 × 3 color correction matrix is used as shown in the equation below.

$$\begin{bmatrix} sR \\ sG \\ sB \end{bmatrix} = \begin{bmatrix} & & \\ & CCM & \\ & & \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

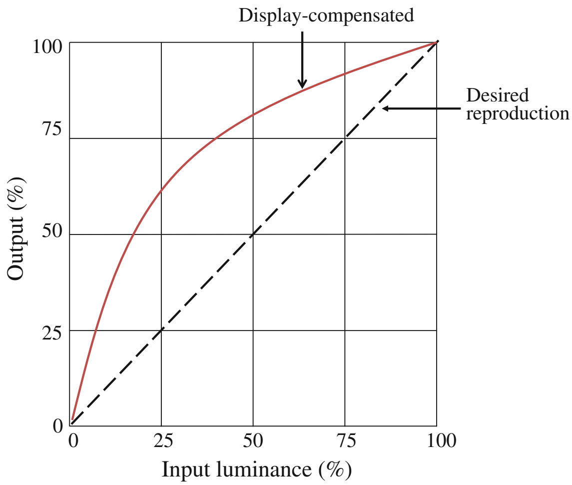
The CCM is constrained to have rows summing to unity (i.e., it is row-stochastic), so that neutrals always map to neutrals. The CCM can be determined in several ways. One approach is to compute response functions such as those in Fig. 8, and then calculate RGB exposure values for many examples of surface reflectances spanning those of natural objects. Because the camera responses are not in general perfect linear combinations of color matching functions, there will not be a single CCM that perfectly maps device RGB to sRGB for all illuminants and reflectance spectra, so some sort of minimization of errors is needed. Alternatively, a physical patch set can be photographed with the camera, and the resulting signal values used as input to a similar analysis.

Terminology used to describe operations related to tone scale is somewhat inconsistent in the literature. We will treat "tone mapping" and "gamma correction" as two separate operations, although they are sometimes combined. Tone mapping is defined to be an intentional manipulation of the tone scale to deviate from accurate tonal rendition, for purposes of producing a more useful, informative, or visually pleasing image. “Gamma correction” then refers just to compensation for the display response, enabling an accurate tonal reproduction over some range, if desired.

In automotive imaging, the most common use of tone mapping is to compress a high dynamic range (HDR) capture into a narrower signal range, so that it can be displayed on a lower dynamic range device and/or processed through an imaging pipeline supporting fewer bits (enabling faster processing and/or lower power consumption). In the latter case, it is advantageous to move the tone mapping operation earlier in the pipeline than shown in Fig. 9, so that the advantage accrues over a larger fraction of the imaging chain. A variety of tone scaling algorithms can be used for HDR compression; a chief distinction is whether they are global (a single tone scale mapping applied to an entire frame) or locally adaptive (varying within a frame).

Gamma correction has historical roots in the first widely available electronic displays, which were cathode ray tubes (CRTs). These devices had a luminance versus driving voltage response that was a power function, with the power, referred to as gamma, often being on the order of two (so that the luminance increased approximately as the square of the voltage). Because there generally was not electronic circuitry present in a CRT to perform transformations on the input signal, it fell to the image pipeline to compensate in advance for this characteristic of the CRT, and this operation was therefore referred to as gamma correction. If the display gamma were near two, then the compensation would be approximately a square root function, as shown in Fig. 10, which assumes a gamma of 2.2 as used in the sRGB standard [3]. As display technology advanced, it became straightforward to incorporate a calibration within the display itself (e.g., in liquid crystal displays), so that no gamma correction would be needed; however, there was no easy way for the industry to make a global transition, as there would always be a mixture of devices requiring and not requiring gamma correction. Consequently, the

Fig. 10 Gamma correction



convention of gamma correction has persisted, and is likely to continue to do so indefinitely.

If the imagery is not displayed, but rather is used for information extraction, as in machine vision applications such as advanced driver assistance systems (ADAS), then gamma correction is not required.

4 Imaging System Measurements

In this section, we will review some of the objective measures that are useful for understanding camera system performance. These include the modulation transfer function (MTF), signal to noise ratio (SNR), and an advanced Fourier metric, noise equivalent quanta (NEQ).

MTF is a measure of signal fidelity as a function of spatial frequency. Signal fidelity is measured by modulation transfer, which is the ratio of the amplitude of an imaged sinusoid to that of the original sinusoid. Spatial frequency can have a variety of units, of which the most common are cycles per mm in the sensor plane, and cycles per pixel (also referred to as cycles per sample). In the latter units, the Nyquist frequency has a value of ½. Signal above this frequency will be aliased to lower, incorrect frequencies, creating image artifacts [4]. Figure 11 shows measured MTF curves of a representative sensor with a pixel size of 3.75 μm, for five wavelengths, up to the Nyquist frequency, which is (1 cycle)/(2*3.75E-3 mm) = 133 cycle/mm. These MTFs are for the sensor only, and do not include a camera lens. Several

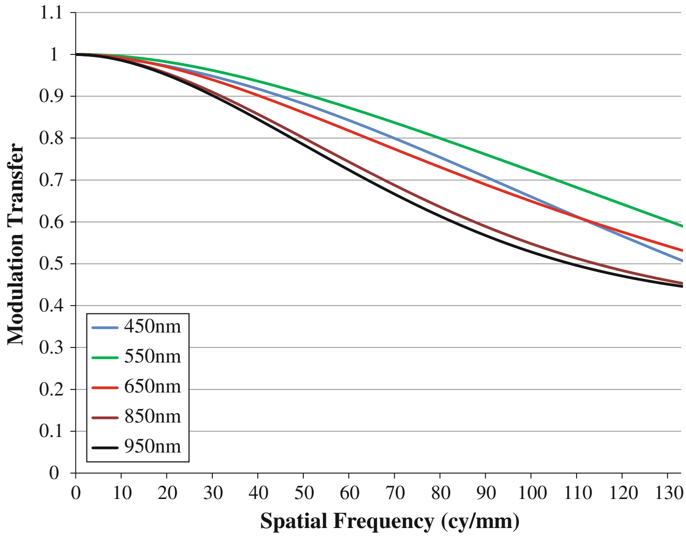


Fig. 11 Sensor modulation transfer function

factors contribute to the wavelength dependence, one of which is that at longer wavelengths, there is less absorption in silicon, so greater lateral spread occurs and the MTF is lower.

Signal to noise ratio (SNR) as commonly measured is the ratio of the mean linear signal in a uniform area, to its standard deviation. It can be quoted either as a simple ratio (e.g., 40:1, which is a good SNR), or in units of decibels (dB) (e.g., $20*\log_{10}(40) = 32$ dB). Decibel units are convenient because, if the noise is well above visual threshold, one dB corresponds approximately to one just noticeable difference (JND) of quality. Figure 12 shows, for a single exposure system at lower signal levels, the SNR in dB versus signal in electrons, on a logarithmic scale (later in this chapter the full SNR curve for a high dynamic range system will be shown). Two limiting cases corresponding to two types of contributing noise are shown. One is the shot noise limit, in which Poisson noise dominates. In this case, the standard deviation is equal to the square root of the mean signal when both are expressed in electrons, so the SNR is also equal to the square root of the signal. The other case is when signal-independent noise sources such as read noise dominate; in aggregate, they define a noise floor. In Fig. 12, the assumed noise floor is 5 electrons. Shot noise and signal-independent noise add in quadrature, so their variances sum to give the overall variance. As shown in Fig. 12, the overall system SNR tends to the noise floor limit at low signal levels, and to the shot noise limit at high signal levels.

In most branches of imaging, SNR is usually quantified as described herein, in a frequency-independent fashion. However, it is possible to quantify noise power as a function of spatial frequency (like MTF), using the concept of noise power spectrum (NPS) [5]. Although NPS falls outside the scope of this chapter, it is

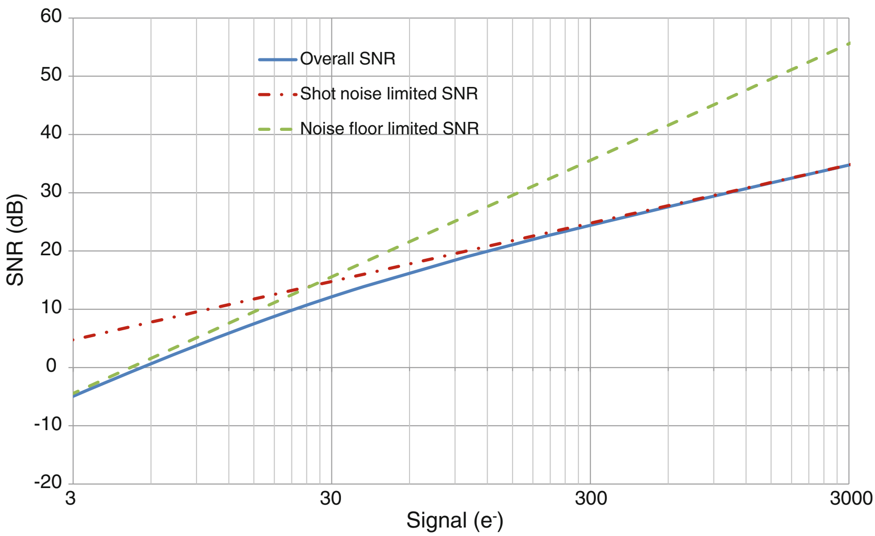


Fig. 12 Signal to noise ratio profile

mentioned because MTF and NPS can be combined to form another Fourier function known as noise equivalent quanta (NEQ). NEQ is the square of the signal to noise ratio as a function of spatial frequency, so higher NEQ is better. It is an exceptionally powerful function, because it can be used to predict task performance via signal detection theory. The volume under the cascade of the NEQ and the frequency power spectrum of the signal to be detected, can be used to predict detection success. NEQ is widely and highly effectively used in medical imaging, where it forms the basis for optimizing medical imaging devices and imaging procedures [6].

As a brief example of the utility of NEQ, consider the detection of a pedestrian by an imaging system and algorithm under difficult conditions. Suppose there were a choice of two sensors having equal silicon area (optical format). Would better performance be expected from a sensor with a smaller number of larger pixels (better SNR in the traditional sense) or a larger number of smaller pixels (having better MTF)? Neither SNR nor MTF alone can be used to answer this question, but NEQ can suffice if the detection algorithm makes effective use of available information. To detect a pedestrian reliably, certain frequencies must be resolved and imaged at sufficient signal to noise. Making the assumption that the highest key frequency is 5 cycles per pedestrian, implies, by the Nyquist criterion, that a minimum of 10 pixels must subtend the height of a pedestrian. Allowing a factor of two margin, we assume that a region of interest containing a potential pedestrian is resampled so that the pedestrian is subtended by 20 pixels.

To construct Fig. 13, NEQ was computed at 5 cycles/pedestrian, over a range of light levels and distances, and the results were plotted as contours of equal $\log_{10}$(NEQ). Performance drops at lower light levels (because of decreased pixel SNR) and at longer distances (because fewer sensor pixels subtend the pedestrian). If an algorithm could reliably detect a pedestrian when the NEQ exceeded about 300, then the 2.5 contour would approximately delineate the region of robust operation. The plot at left depicts the results for a 1-megapixel sensor, and the plot

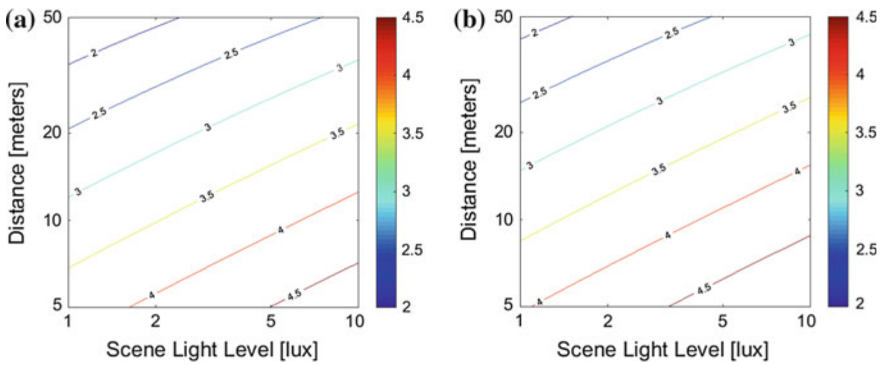


Fig. 13 **a** Log_{10}(NEQ) contours of 1-megapixel sensor. **b** Log_{10}(NEQ) contours of 2-megapixel sensor

at right for a 2-megapixel sensor having equal optical format (silicon area) and using the same camera lens. The 2-megapixel sensor maintains slightly higher NEQ at longer distances and lower light levels, as shown by how closely the 2.5 contour approaches the top left corner of the plot. The 2-megapixel sensor performs better because, at a given distance, a larger number of pixels subtend the pedestrian. However, this result might not have held if the pixels in the two sensors were considerably smaller, because the SNR performance of the even smaller pixels in the 2-megapixel sensor might have been poorer due to proportionally larger noise floor effects.

5 High Dynamic Range Imaging

In this section, we consider a challenge of particular importance in automotive imaging: the need to capture extremely high dynamic range scenes (greater than about 100 dB).

An image sensor can increase the effective dynamic range beyond that of an individual pixel by utilizing multiple exposures. There are basically two methods by which this can be achieved; temporal or spatial. Image sensors of both types are available in the automotive market. Here we describe the principle of dynamic range extension, and the basic operations as well as the benefits and drawbacks of each of these methods. Actual implementations usually include additional post-processing and artifact-reduction techniques available in the sensor or in a separate device.

Figure 14 shows the SNR profile diagrammatically for a temporal three-exposure HDR system. The x-axis is proportional to scene luminance, and the y-axis is proportional to linear SNR. The SNR segments are curved principally because shot noise SNR is proportional to the square root of the signal. Proceeding from left to right (increasing scene luminance), the longest exposure, denoted E1 and shown in red, provides the best SNR until the pixel saturates or its response begins to become non-linear. When this occurs, information must instead be drawn from the intermediate exposure (E2, yellow), again until the pixel saturates. There is an SNR drop (discontinuity) at the point where there is a transition from use of E1 to E2. This can sometimes be seen in parts of a scene where there is a slowly varying sweep, but it can be mitigated by digitally mixing in some E2 signal as E1 approaches saturation, to round off the transition. When E2 approaches saturation, there is a transition to the shortest exposure, E3. The dynamic range of the entire system is determined by the ratio of the scene luminance saturating E3 (right edge of plot) to the scene luminance producing a threshold SNR value, usually assumed to be unity (near the left edge of the plot). This system dynamic range exceeds the dynamic range of a single pixel (usually approximated as the ratio of the linear full well to the noise floor) by approximately the ratio of the longest exposure to the shortest exposure. Although this ratio could be made almost arbitrarily large, if it is too large, the SNR drops at the transitions are too great and the SNR values at the bottom of the

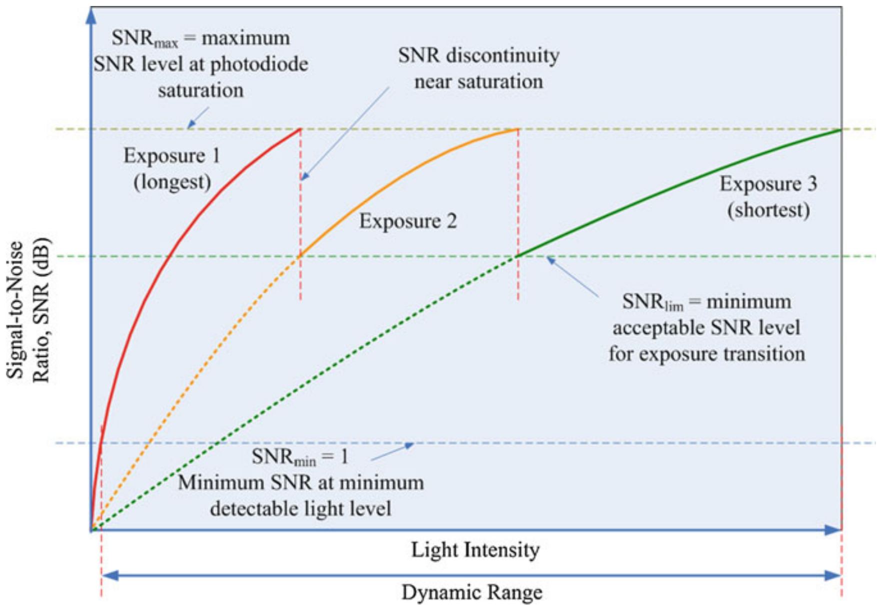


Fig. 14 SNR profile for multi-exposure HDR

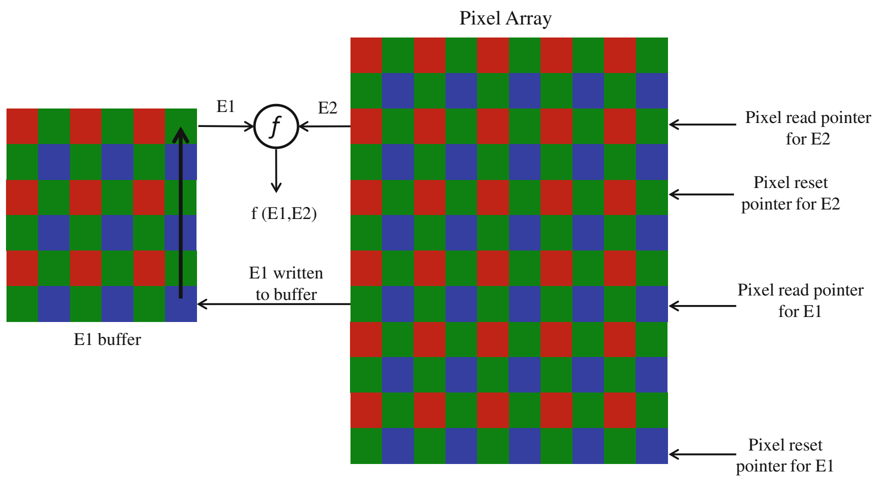


Fig. 15 Temporal HDR

transitions are too low. Adding a fourth exposure of course allows additional dynamic range extension.

Figure 15 shows the readout process for a simpler two-exposure temporal system. The two exposure address pointers are named E1 (long exposure) and E2 (short exposure). During readout, the pixel row at address E1 is read and the E1 pixel data stored in the line buffer memory. The line buffer memory provides temporary storage so that both exposures for each pixel may be aligned and combined to form a linear signal. Once the E2 read pointer is aligned with the data that was read out during the E1 operation, the two exposures, E1 and E2 can be combined to create the final output signal. This scheme requires two pixel readouts and ADC conversions per pixel in the array so the overall sensor array readout circuits must operate at higher speed and consequently higher power.

Typical systems might support ratios of 4, 8 and 16. The greater the ratio supported, the larger the memory required to accommodate the readout so that each row can subsequently be recombined. The timing of these systems can become very complex to ensure that the reset and sample pointers of each exposure do not overlap and cause data corruption.

Depending upon the technique used to store and combine the multiple exposures, and the number of exposures used, the temporal technique can lead to artefacts due to motion in the scene between the exposures. This is typically countered with post processing techniques but if these are turned off, artefacts such as that exhibited in Fig. 16 become evident. In this figure, there is ordinary motion blur due to high angular motion during the exposure, but there are also color artefacts behind the headlamp. These arise from different color channels being selected from different exposures because of different saturation points. There are many motion compensation algorithms that can be used to mitigate this type of artefact.

Spatial HDR can be implemented in many ways. One common method is to interlace the exposures as shown in Fig. 17 where alternate row pairs pertain to either E1 or E2. The trade-off of interlacing the exposures is that fewer pixels are available for each exposure image and can affect the overall captured image resolution in the vertical direction. Figure 17 shows a Bayer pixel array where pairs of rows alternate exposures E1 and E2 such that odd row pairs use E1 exposure and even row pairs use E2 exposure. It also shows how separate row address pointers for pixel reset are used to start the exposure times for E1 and E2 such that the readout address pointer is aligned to readout each row with the desired E1 or E2 exposure. Since the E1 and E2 exposures are interleaved, the readout of the pixels can be performed the same as a standard sensor without requiring large memory buffers for readout alignment.

Fig. 16 Temporal HDR artefacts

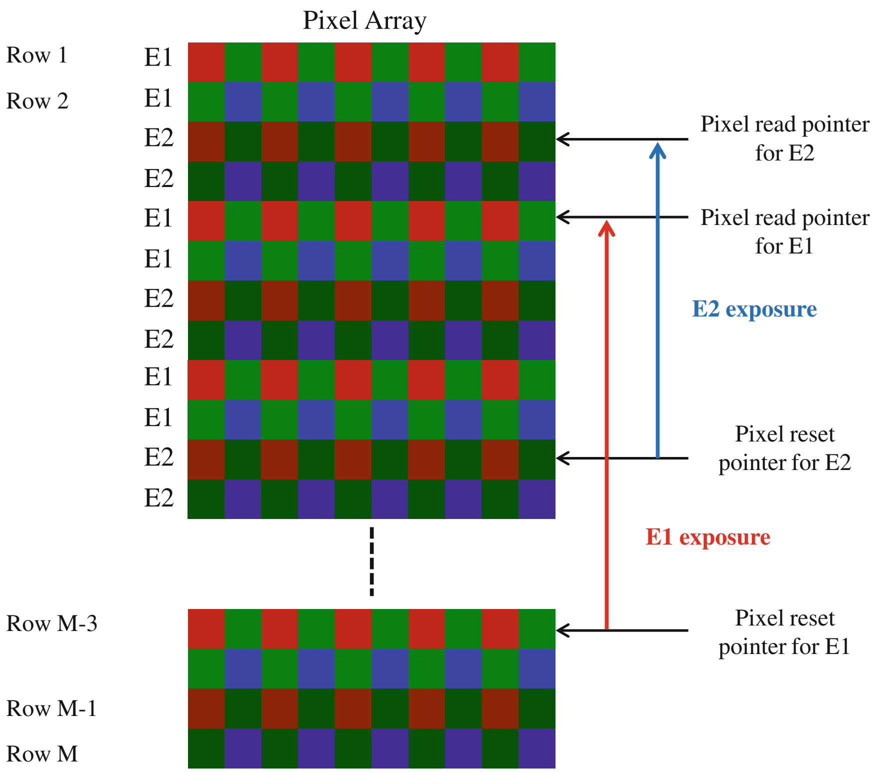
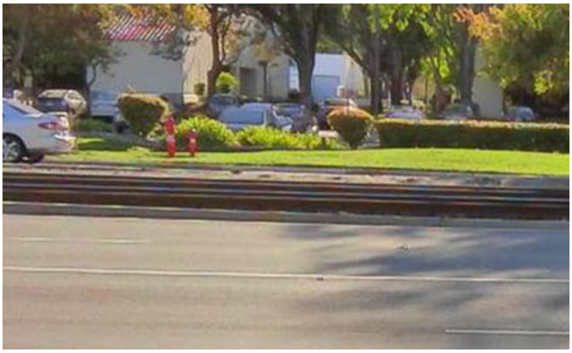


Fig. 17 Spatial HDR

Spatial HDR can also exhibit artifacts that can be fixed with post-processing. With the row-interleaved technique shown above, one common type of artifact is seen on diagonal edges. Due to the reduced vertical resolution available in any one exposure, there is a propensity for color aliasing artifacts as seen in Fig. 18 on the rooftop and on the rail track in the center of the image.

Fig. 18 Spatial HDR artefacts



6 Light Source Flicker

A significant challenge for imaging systems in general, and HDR systems in particular, is avoidance of flicker when capturing scenes illuminated by, or containing, time-varying light sources. Consider first the simplest case of a single-exposure system imaging under a light source that is modulated by the power line frequency. For example, a fluorescent tube powered by a 50 Hz line will typically vary in output at a frequency of 100 Hz, corresponding to a period of 10 ms. In a rolling shutter exposure, different sensor array rows will be captured at different times and so will be exposed by differing amounts as they sample different periods in the output oscillation. Figure 19 shows an example of an image with such 50 Hz flicker.

Given a knowledge of the line frequency in a geographical region, or inferring the same from captured image analysis, permits a solution to this problem under many circumstances. If the integration time is set to a multiple of the light source period (10 ms in this case), then each row samples an integer number of cycles and the effect is averaged out. Figure 20 shows a timing diagram for such a case. At the top of the figure, the 50 Hz line frequency is shown as a sinusoid, which is effectively rectified by the fluorescent ballast to produce a 100 Hz output oscillation. The integration of the sensor pixels in one row is shown at the bottom of the figure. Because the integration period of 20 ms is a multiple of the output oscillation period of 10 ms, the row collects an integer number of output cycles (two), and so the signal is independent of the phase of the row integration period relative to the phase of the output cycle, and flicker is eliminated.

In a temporal HDR system, the above approach may work well for the longest exposure; however, the intermediate or short exposures may be too short to be a multiple of the power cycle period. Thus, although flicker may be avoided in E1, there may be E2 and E3 flicker. In addition to this problem, in recent years, light emitting diodes (LEDs) have become pervasive in automotive lighting (e.g., headlamps and brake lights), traffic control (e.g., road signs and traffic lights), and general illumination. LEDs are advantageous in many respects, including

Fig. 19 50 Hz flicker

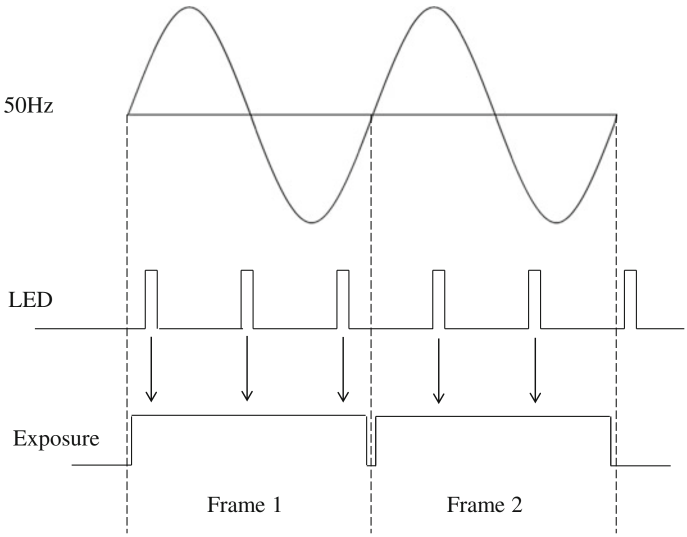


Fig. 20 Flicker timing diagram

consuming less power at equal light output, and having longer lifetimes. However, they pose serious imaging challenges, because they operate over a very wide range of frequencies (up to about 500 Hz) and can have low duty cycles (down to about 10 %), producing trains of narrow pulses that can be widely spaced in time. For example, the middle of Fig. 20 shows an LED pulsing at 125 Hz, with a duty cycle of 12.5 % (so it is on for 1 ms during each 8 ms period). During the integration time of the first video frame, three LED pulses are recorded, but during the second video frame integration, only two pulses are recorded, which would lead to temporal flicker in the video stream.

Not only is their frequency variable and generally unknown, but LEDs within a device may operate with different phases. For example, an LED sign may drive the

Fig. 21 LED flicker

LEDs in different positions with different phases, so that power dissipation is evened out rather than concentrated in time. This effect is shown in Fig. 21. Only some LEDs in the speed limit sign on the right side of the tunnel are captured successfully during the integration time; other LEDs pulse on at times when the pixels onto which they are imaged are not integrating. In combination, variable frequency and variable phase make LED flicker a very challenging problem.

7 Representative System Solution

A system solution requires that the image sensor is first packaged, such that it is sealed against humidity and physical contamination. This packaged device is then integrated into a camera head and finally this is connected to a control unit and a display.

Image sensors for automotive applications require unique packaging solutions. Here we present two popular package types, the imager ball grid array (iBGA) and the imager chip scale package (iCSP). The iBGA package shown in Fig. 22 consists of a substrate to which the image sensor is attached using epoxy or film. Gold wires are used to connect the die pads on the image sensor to bond fingers on the substrate using a ball wire bonder. The glass lid covers the image sensor and the wires are protected by the glass lid or mold encapsulate. Finally, at the bottom side of the substrate is routing from the bond fingers to an array of solder balls.

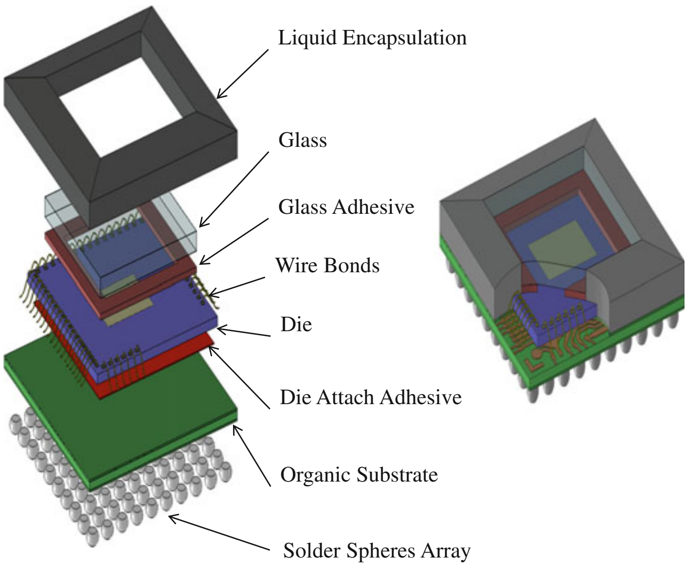


Fig. 22 Imaging ball grid array package

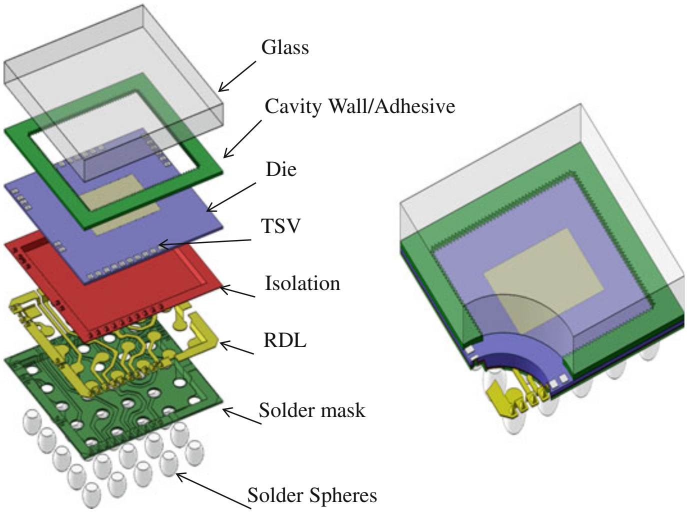


Fig. 23 Imaging chip-scale package

The iCSP is assembled at the wafer level. In a typical construction shown in Fig. 23, a sheet of glass the same dimension as the silicon wafer is patterned with a dam of epoxy that mirrors the outline of each of the die on the wafer. This is placed over, and bonded to the wafer. Vias are formed through the back of the image sensor substrate to reveal the underside of the specially designed connection pads. An isolation layer is placed onto the back of the die to isolate it prior to patterning an interconnect or redistribution layer (RDL), which is used to route connections from the silicon bond pads to the solder balls.

A typical CMS will integrate the packaged image sensor into a camera head unit and connect this to an electronic control unit and a display as shown in Fig. 24. Depending upon the architecture deployed the electronic control unit (ECU) and the display may be co-located. It is anticipated that the cameras in a CMS will have a higher resolution than that of the display to allow for panning (the equivalent of moving the mirror for certain maneuvers) and to permit other automated functions such as blind spot detection. Today image sensors designed for automotive

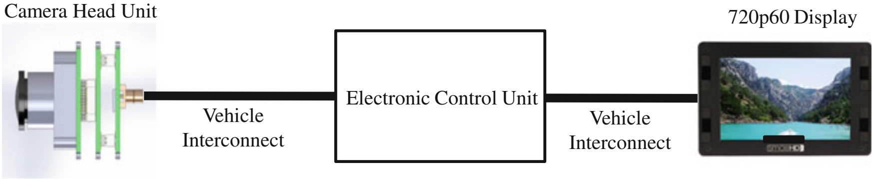


Fig. 24 Vehicular CMS

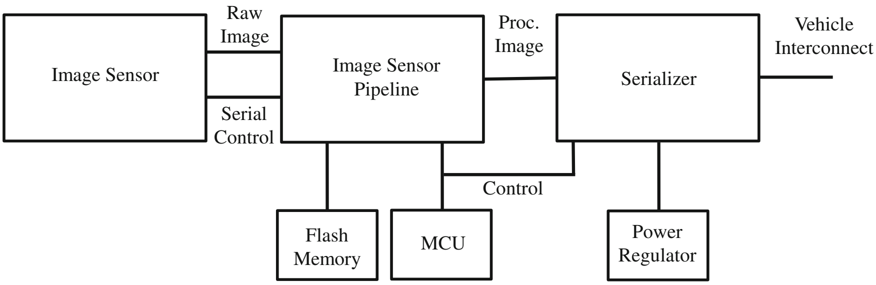


Fig. 25 Camera head schematic

applications are available in approximately 0.3 megapixel, 1 megapixel and 2 megapixel resolutions and while these lower resolutions will likely be utilized in early CMS, it is anticipated that the desire for increased display quality will lead to the requirement for higher resolution cameras. Increasing camera resolution and frame rates is a complex system decision as it can lead to increased power consumption and higher bandwidth requirements on the vehicle interconnect. For example if we consider a current 2 megapixel sensor and if each pixel is represented by 12 bits (having integer values between zero and 4095) and it operates at 60 frames per second, then this will require approximately 2 gigabits per second of interconnect bandwidth to the ECU. Today the vehicle interconnects for a CMS could be implemented with a low voltage differential signaling (LVDS) system, which can offer up to 3 Gigabits per second of bandwidth. Increases in resolution or frame rate will need to be supported by commensurate increases in connectivity bandwidth.

An exemplary camera head unit is shown in Fig. 25. The camera head unit contains the image sensor, a micro controller unit (MCU), Flash memory, and power regulator, as well as a device to serialize the image data for transmission over the vehicle interconnect. The MCU and Flash memory are used to configure the camera head components at startup with basic parameters such as resolution and timing. The vehicle interconnect systems deployed for CMS can be bidirectional, transmitting image data from the camera head, and control information from the ECU. Some interconnects may also supply power or there may be a dedicated power supply connection to the camera head.

8 Safety

The complete CMS will possibly require an ISO 26262 Automotive Safety Integrity Level (ASIL) rating [7]. An ASIL rating is assigned to the complete system, and using a process known as ASIL decomposition, each subcomponent of the system (such as the camera) will receive a designated ASIL rating. Individual components may not be required to have their own ASIL rating, but may support the system design through the use of watchdog timers and other signal monitoring techniques. The image sensor will, by definition of ISO 26262, be classified as a “safety element designed out of context” due to the fact that it will normally be designed prior to the functional safety concept of the system being completed.

The primary failure mode that may be of concern in a CMS is that of a frozen image from an earlier video frame on the screen, as this would not be immediately obvious to a driver. Failure modes of this nature would require a complete frame of memory to exist, and for the write mechanism of that memory to fail. The typical image sensor utilizes a destructive read of the pixel array, and so cannot retain a complete image frame. The image processing techniques identified in this chapter can be implemented without a complete image frame of memory, although some systems may utilize a complete frame of memory. Image sensors can include a variety of additional mechanisms to assist a system processor in identifying a failure. Some typical mechanisms on new image sensors, specifically designed for automotive applications, include monitors that ensure all rows and columns are reading out successfully and in the correct sequence on each frame. In addition, the sensors are protected by error correcting codes (ECC) on memory contents and cyclic redundancy checks (CRC) on interfaces.

9 Conclusion

CMS is an exciting new application, with considerable opportunity for innovation, particularly in the image sensor and processing components. Areas where technological progress is anticipated in the sensor include improved flicker mitigation, higher dynamic range, reduced dark current, and lower noise. Image processing algorithms will become more sophisticated, providing higher image quality and enabling new features based upon machine vision. Future system specifications will likely require higher resolution (above 2 megapixels), lower power, and higher frame rate (60 frames per second or greater). These advances in image sensors and processing will contribute to the greater adoption of CMS systems in automobiles in the coming years.

References

1. IEC 61966: Multimedia systems and equipment—colour measurement and management—Part 2–1: colour management—default RGB colour space—sRGB. Amendment 1, International Electrotechnical Commission, Geneva (2003)
2. Keelan, B.W.: Handbook of Image Quality: Characterization and Prediction. Marcel Dekker, Inc., New York (2002) (Sect. 18.2)
3. Dainty, J.C., Shaw, R.: Image Science. Academic Press, London (1974) (Ch. 8)

4. International Commission on Radiation Units and Measurements, Report 54: Medical imaging —the assessment of image quality. Bethesda (1996)
5. Takayanagi, I.: CMOS image Sensors (Ch. 5). In: Nakamura, J. (ed.) Image Sensors and Signal Processing for Digital Still Cameras. Taylor and Francis, New York (2006)
6. Velichko, S., Hynecek, J., Johnson, S., Lenchenkov, V., Komori, H., Lee, H., Chen, F.: CMOS global shutter charge storage pixels with improved performance. IEEE Trans. Electron Devices, to be published, 2015
7. ISO 26262: Road vehicles—functional safety, parts 1–9 (2011) and Part 10 (2012). International Standards Organization, Switzerland

Optical Effects in Camera Monitor Systems

Christian Faber and Patrick Heinemann

Abstract This study deals with the optical properties of camera-monitor-systems (*CMS*) used in road vehicles. The goal is to specify the requirements for such systems in a way that the human perception of the presented scene does not differ in any critical way from the perception obtained by direct view or via classical mirror solutions. To arrive at objective and verifiable requirement specifications, it is important to define measurable physical quantities that completely determine the system behavior with respect to the resulting optical stimulus. As a prerequisite, first a short review of the basic photometric and radiometric measures and their interrelations is presented. Furthermore, the crucial quantities for the given application are identified, and typical values are given that are needed for the practical assessment of a given CMS. Second, image artifacts induced by spurious reflections, scattering and other shortcomings of the lens system that are related to stray light are examined. This is particularly important for the application of CMS in road vehicles, as these camera systems have to be able to render poorly lit scenes and at the same time cope with the intense headlights of trailing or approaching cars. Finally, in order to assess the performance of a given CMS in this respect, corresponding measurement setups and test procedures are proposed.

Keywords Image artifacts · Stray light · Glare · Flares · Aperture ghosts · Test setups for CMS

C. Faber (✉)
Sensor Technology and Image Processing, University of Applied Sciences Landshut, Landshut, Germany
e-mail: christian.faber@haw-landshut.de

P. Heinemann
Pre-Development Driver Assistance Systems, Audi Electronics Venture GmbH, Gaimersheim, Germany
e-mail: patrick.heinemann@audi.de

A. Terzis (ed.), *Handbook of Camera Monitor Systems*,
Augmented Vision and Reality 5, DOI 10.1007/978-3-319-29611-1_6

1 Introduction and Groundwork

Before the actual analysis of properties and optical effects of camera-monitor-systems used in automotive environments can start, some groundwork has to be provided to set the stage and to introduce the notation used in the subsequent sections. For this, some basic concepts of photometry and radiometry are revisited particularly with regard to the given application.

1.1 Photometry Versus Radiometry

The optical stimulus forwarded to the brain providing the input for the driver's decisions and (re-)actions is a direct function of the distribution of electromagnetic radiation impinging on the retina. As this radiation is sampled by different kinds of photoreceptor cells (color-sensitive cone cells and purely intensity-sensitive rod cells), it seems natural to define physical quantities that incorporate the physiological weighing introduced by these sensor elements. This is achieved by defining a standardized spectral sensitivity curve for the human eye (the so-called *V(λ)-curve*, also called *luminous efficiency function* or *luminosity function*). For every radiometric quantity F_e (such as *radiant flux* Φ_e, *radiant intensity* I_e, *radiance* L_e, *irradiance* E_e etc.—see Sect. 1.2), the corresponding photometric quantity F is defined by weighting the spectral components of F_e according to the *V(λ)-curve* ([1, 2]):

$$F \equiv K_m \cdot \int_0^\infty \frac{dF_e}{d\lambda} \cdot V(\lambda) \cdot d\lambda;\ K_m \equiv 683{,}002\,\frac{\mathrm{lm}}{\mathrm{W}} \tag{1}$$

The quantity F_e can be any of the above radiometric quantities Φ_e, I_e, L_e, E_e. The name of the corresponding photometric quantity is obtained by replacing the term "radiant" with the term "luminous" (e.g. *luminous flux* Φ, *luminous intensity* I, *luminance* L, *illuminance* E etc.). The crucial criterion for the assessment of a CMS is the final impact on the (human) driver. Therefore, the physical quantities to be considered are the <u>*photometric*</u> ones. However, there are significant differences between the luminosity functions for *photopic vision* (high intensity, daylight) and *scotopic vision* (low intensity, night scenes)—see Fig. 1.

When defining limit values for test scenarios, this has to be taken into account. Especially for the assessment of glare-scenarios at night, the scotopic luminosity function $V'(\lambda)$ should be applied.

In this context, it should be noted that for the CMS case definition '*night condition*' (see [4]), it may be advisable to define certain standards regarding the display properties. To avoid bedazzlement of the driver looking with his (scotopically adapted!) eyes at the display, for one thing, luminance limits

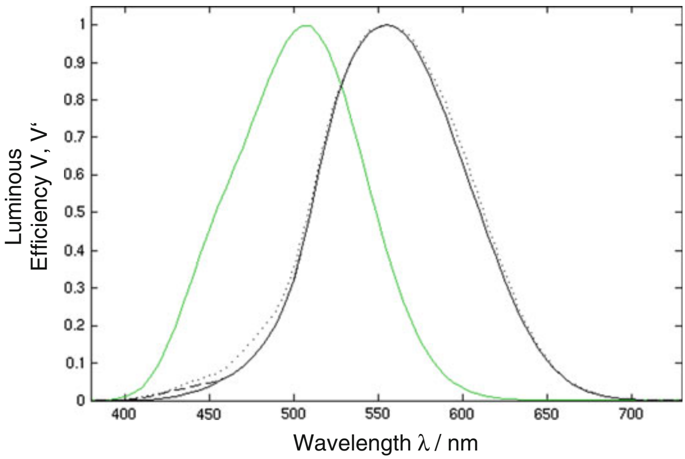


Fig. 1 Luminosity functions for photopic vision ($V(\lambda)$, *black curve*) and scotopic vision ($V'(\lambda)$, *green curve*). Image from [3]

(see Sect. 1.4) for the display should be defined for this situation. To go even further, it could be helpful to demand that the spectral composition of a night scene presented on the display should also be adapted to the scotopic sensitivity curve, e.g. by making use of shorter (blue) wavelengths. However, in this context also psychological aspects have to be considered—it may even be advisable to completely *abandon* the use of different colors when displaying night scenes altogether, or at least to refrain from using large hue ranges in this case, as otherwise this may have negative effects on accommodation.

1.2 Review of Photometric Quantities

The ultimate measure determining the driver's sensory impression is the amount of luminous (i.e. physiologically spectrally weighted) energy impinging on the retina per unit time and unit area. Luminous energy per unit time is called *luminous power* or *luminous flux* Φ. The unit of Φ is the *Lumen*; according to the standard (maximum) luminous efficacy[1] of $K_m = 683$ lm/W (see Eq. 1), a radiant power of 1 W at a wavelength of 555 nm corresponds to a luminous power of 683 lm for photopic vision (see [2]).

[1]Mind that this definition has been changed in 1979 by the CGPM (Conférence Générale des Poids et Mesures)!

Thus, the relevant quantity at the end of the imaging chain is the luminous flux Φ per unit area A on the sensing plane (retina). This quantity is called *illuminance*:

$$E \equiv \frac{d\Phi}{dA} \tag{2}$$

The unit is lux; 1 lux = 1 lm/m^2. The illuminance distribution in the image plane is determined by the properties of the imaging system (classically the combination mirror—eye) and by the observed scene itself. To identify (and quantify) which aspect of the scene is actually governing the resultant signal, it is helpful to define additional densities of the luminous flux—not only with respect to area, but also with respect to solid angle (and a combination of both).

The density of luminous flux with respect to solid angle is called *luminous intensity I*:

$$I \equiv \frac{d\Phi}{d\Omega} \tag{3}$$

The unit is candela; 1 cd = 1 lm/sr. The luminous intensity can be used to characterize a pointlike light source. For a spherical (i.e. isotropic) radiator, the emitted radiation (resp. luminous flux) is constant in all directions. In this case, the specification of one value given in candela is sufficient. However, most light sources show a prominent angular variation of their luminous intensity, as additional optics like collimators are used to concentrate the luminous flux into a desired direction. In this case, the whole function $I(\varphi, \theta)$ has to be specified. The candela value stated in those circumstances commonly refers to the luminous intensity in the direction of maximum emission. The function itself can either be given as a measurement plot or be represented by some sort of model curve.

If a (self-luminous) light source or an (externally illuminated) object is not pointlike, it makes sense to additionally relate the luminous intensity to the unit area of the source in a given direction. More precisely, to eliminate 'trivial' geometrical projection effects, this areal density is taken with respect to the projected area element perpendicular to the viewing direction. The resulting quantity is called *luminance L*:

$$L \equiv \frac{d^2\Phi}{d\Omega \cdot dA \cdot \cos\theta} \tag{4}$$

The variable θ denotes the angle between the surface normal and the specified direction. The unit of luminance is candela per square meter. Typically, the luminance of a spatially extended surface is a function of position (x, y) as well as direction (φ, θ). Therefore, quite similar to the case of luminous intensity, for a complete characterization of a given surface also the angular distribution of the luminance has to be given. For a large number of surfaces, it is a very good approximation to assume that the luminance is independent of the angle. These surfaces are called *Lambertian*; they appear with a constant perceived brightness when observed from different directions with an imaging system. According to the above definition, their luminous intensity has a theta—dependence of $I(\theta) = I_0 \cos\theta$ (as opposed to spherical scatterers resp. isotropic radiators with $I(\theta) = const$). For most non-specular 'natural' objects and surfaces observed in a road scene (pavement, houses, trees, the sky, …), the Lambertian model is an extremely good approximation. However, when considering glare effects by headlights of approaching or trailing vehicles, the angular distribution of the light source is decisively non-Lambertian (and in most cases also non-spherical). When defining test cases for these situations, this has to be taken into account (→ angular adjustment of the test setup required (!); see Sects. 2.3, 3.2, and 3.4).

Finally, if the illuminance is not considered as a luminous flux impinging on a passive (sensing) area, but as the flux emitted by an active (luminous) surface, this quantity is called *luminous emittance M*. Such a light-emitting surface can either be self-luminous or reflecting/scattering the light coming from an external source. The luminous emittance has the same units (lux) and characteristics as the illuminance.

The interrelations of the photometric quantities introduced above are summarized in Fig. 2.

Table 1 gives a comprehensive overview of the introduced quantities and their units.

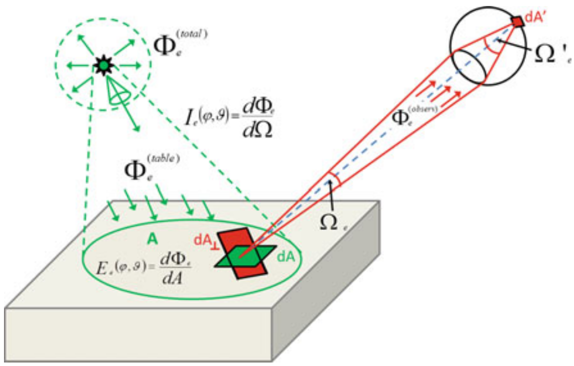


Fig. 2 Interrelation between photometric quantities, illustrated by the 'canonical' case of an imaging system (e.g. eye) observing an illuminated surface (i.e. tabletop)

Table 1 Comparison of radiometric and photometric quantities

Radiometric quantity (physics)	Photometric quantity (physiology)
Radiant flux Φ_e [W]	Luminous flux Φ [lm]
Irradiance E_e [W/m^2]	Illuminance E [lm/m^2]
Radiant intensity I_e [W/sr]	Luminous intensity I [cd = lm/sr]
Radiance L_e [W/(sr·m^2)]	Luminance L [cd/m^2]

1.3 Notation Conventions

When describing imaging systems, conceptually it is extremely helpful to distinguish between quantities defined on the object- and on the image side. Therefore, it is a widely used convention to denote quantities in object space with the letters defined in Sect. 1.2 (Φ, E, L, Ω, ...), whereas the corresponding image-side quantities are denoted with a prime (Φ', E', L', Ω', ...). Using this convention allows for an extremely handy short-hand-notation (e.g. $L = L', \Omega'/\Omega \sim \beta^2, E' = L \cdot \Omega'$ etc.).

The ratio between the (lateral) size of the image and the original object size is called *image scale* or *image ratio* and is denoted by the greek letter β. Aperture angles are denoted by u (object side) respectively u' (image side). Solid angles are denoted by Ω. The solid angle element is $\mathrm{d}\Omega = \sin\theta \, \mathrm{d}\theta \, \mathrm{d}\varphi$, $\int \mathrm{d}\Omega = 4\pi$.

1.4 Luminance as the Crucial Quantity

Using the above definitions, it becomes apparent by what property of the scene the final image illuminance is determined: The quantity to consider is the luminance distribution $L(x, y, z; \varphi_O(x, y, z), \theta_O(x, y, z))$ in object space, where each single value has to be taken for the direction $\varphi_O(x, y, z)$, $\theta_O(x, y, z)$ pointing from the surface element (located at x, y, z) to the imaging system. If all surfaces contributing to the image can be regarded as Lambertian, the direction dependence can be dropped. In this case, it is sufficient to consider the mere spatial distribution of the luminance $L(x, y, z)$. Furthermore, if the interrelation between the 3-dimensional world coordinates (x, y, z) and 2-dimensional image coordinates (x_i, y_i) is known—for example by use of an appropriate camera model providing functions $x_i(x, y, z)$, $y_i(x, y, z)$—it is sufficient to consider a corresponding 2-dimensional luminance distribution $\mathrm{L}(\mathrm{x_i}, \mathrm{y_i})$ denoting the luminance value of the surface element in object space that is imaged at the image coordinates (x_i, y_i). However, if the assumption of Lambertian surfaces is not valid, care has to be taken that these values are actually evaluated for the proper directions.

Table 2 Typical luminance values relevant for camera-monitor-systems in road vehicles

Object	Luminance (cd/m^2)
Clear sky	8000
Night sky at full moon	0.1
60 W light bulb	120×10^3
TFT white	500
TFT black	0.8
White LED	50×10^6
Sun on a clear day at noon	1.6×10^9

Table 2 shows some typical luminance values that are relevant for the case of a camera-monitor-system observing a road scene. Please note the enormous dynamic range of the different values involved. The human visual system copes with these demanding requirements by a clever combination of different adaptation mechanisms. For a CMS, the use of a High Dynamic Range camera (HDR) is advisable. However, it won't be avoidable that for certain constellations (direct observation of the sun or headlights of a trailing vehicle directly shining into the system) the system will be driven into some sort of saturation.

From the explanations presented above, it should be clear why the luminance is the quantity of choice to characterize the photometric properties of the object to be imaged. This is the reason why luminance is sometimes also called 'photometric brightness' or just 'brightness'. Things are especially easy when only Lambertian surfaces are present in the observed scene. In this case, the perceived brightness does not depend on the distance z of the object. The reason is that for an approaching object, there are two effects that exactly cancel each other out: At first, for a standard (endocentric) imaging system, a closer object is imaged with a larger image ratio. Therefore, for the moment one would expect a *reduction* of image plane illuminance as the light is distributed over a larger image area in the sensor plane. However, at the same time, for each single image point *more light* (luminous flux Φ) is collected by the system as the effective aperture angle Ω increases with decreasing object distance (if the aperture stop resp. entrance pupil of the system remains constant!). The corresponding *gain* of image plane illuminance due to this second effect exactly compensates the loss induced by the first one. This is actually the reason why the 2-dimensional trapezoidal shape on a display representing a street that runs in 3d object space from the observer's position to the horizon typically has a constant gray value. In other words, for Lambertian surfaces the perceived brightness provides no cue for depth recognition.

It should be noted that this observation is only true if the aperture of the system (solid angle Ω) is always completely filled by the impinging light. This may not be the case for perfectly collimated light sources (like lasers). However, for typical camera systems used in road vehicles, the size of the entrance pupil (i.e. the 'effective' aperture area determining the solid angle Ω) is rather small, so that the possibility of an only partially filled pupil can be neglected for all practical purposes.

A further 'nice' property of Lambertian surfaces is the fact that if a passive, i.e. not self-luminous surface is illuminated by an illuminance E, the reflected luminance L_r (sometimes called 'borrowed luminance') can be calculated in a straightforward way:

$$L_r = \frac{E \cdot \rho}{\pi} \tag{5}$$

Here, ρ denotes the local reflectivity of the surface. For a perfectly white (not absorbing) surface, $\rho = 1$.

Finally, a remarkable property of the luminance is the fact that it is an invariant under imaging processes (if there are no absorption losses in the imaging system):

$$L = L' \tag{6}$$

Therefore, when asking for the illuminance distribution E' on the image plane, this can easily be calculated by

$$E' = L \cdot \Omega' \tag{7}$$

Here, Ω' denotes the solid angle spanned by the aperture on the image side. More precisely, the above equation is only correct for a Lambertian (constant) luminance distribution L. For the non-Lambertian case, the right side has to be replaced by $\int L(\Omega')\, \mathrm{d}\Omega'$.

1.5 Purpose of a Camera-Monitor-System

As a straightforward application of the observations presented above, one can wonder what luminance should be chosen for the display of a given CMS system in order to ensure exactly the same sensory impression for the observer, i.e. exactly the same illuminance distribution on his retina as if he were observing the scene directly or via a classical mirror.

As described in the previous sections, the task of an (ideal) imaging device is to generate an *illuminance* distribution $\boldsymbol{E'}$ on the sensing area that is proportional to the *luminance* distribution $\boldsymbol{L}$ in object space. The proportionality constant depends on the properties of the imaging system (more precisely on the solid angle Ω' of the image-side-aperture) and is determined by the so-called "etendue" of the system. In this context it should be noted that if the imaging system properties are changed between different application situations (e.g. by using a dynamically adjustable aperture), this has to be taken into account when defining the appropriate test cases!

Now, one approach could be to require that the luminance of the scene presented on the display should be exactly equal to the original luminance in object space. At least for Lambertian objects, this would result in the same illuminance of the driver's retina as if he were observing the scene via a mirror. However, this is not feasible in practice, as the typical luminance values of the original objects span such an enormous dynamical range that this can never be covered by a display (even if the corresponding values could be recorded on the input side, e.g. by using a high dynamic range camera). Furthermore, even if this were possible, it wouldn't always be advisable to faithfully reproduce the original luminance distribution, as for example it is not desired to preserve the glare of a headlight of a trailing vehicle on the display. In addition, one advantage of using a CMS is to be able to enhance the visual representation of the scene (e.g. by augmented reality). However, it should be noted that for night scenes, it will be critical to provide a sufficiently *low* display luminance, which is not that trivial technically. This is another point which should be tested when assessing a given CMS.

For a real imaging system, the illuminance distribution on the sensing area (retina/chip) will always deviate from the original luminance distribution. This can be caused by imaging aberrations, but also by glare effects like ghosts and flares (see Sect. 2). Every effect that can cause an illuminance value at a given image location where there is no corresponding luminance signal in object space has to be considered as a possible (optical) malfunction of the imaging process, i.e. a system flaw, and is to be specified accordingly.

1.6 Recap—Summary of the Fundamentals

The goal of the preceding subsections was to set the stage for the subsequent considerations by reviewing the basic physical quantities required and identifying those quantities and aspects that are crucial for the given application of using a camera-monitor-system in road vehicles. Particularly, the physical quantities and standard notation needed for the definition of system requirements as well as test setups have been introduced and explained. The most important points can be summarized as follows:

- The crucial quantity in object space defining perceived image brightness is the luminance [cd/m^2].
- As the spectral sensitivity of the human visual system depends noticeably on the scene brightness, it is advisable to use the scotopic luminosity function $V'(\lambda)$ for night driving scenarios.
- To define the optical property of the scene, it is not sufficient to stipulate only one scalar luminance value per source point—also the angular distribution of the (possibly reflected) luminance has to be specified (e.g. lambertian [L = const], isotropic [$L \sim \cos^{-1}\theta$] or collimated [$L \sim \delta(\theta)$]). This is particularly important for the definition of test setups, as for example certain illuminance values required for a realistic simulation of glare scenarios can only be achieved with collimated sources.
- In an imaging process, luminance is an invariant quantity. Therefore, the image plane illuminance E' responsible for the sensory impression on the retina and the resultant grey value on the camera chip can easily be calculated by $E' = L \cdot \Omega'$ (Lambertian case).
- In the case of Lambertian surfaces, the perceived brightness is independent of object distance.
- If a (passive, not self-luminous) Lambertian surface is illuminated with an illuminance E, the reflected luminance can easily be calculated by $L_r = E \cdot \rho/\pi$, where ρ denotes the local reflectivity of the surface.

2 Stray Light Artifacts

Now, as the stage has been set and the necessary tools have been provided, it is time to (literally) shed some light on the possible image artifacts induced by spurious reflections, scattering and other shortcomings of the lens system that are related to stray light. The goal is to explain the origin of these effects, classify different types of stray light artifacts (so-called ‘ghosts and flares’), and to give clues as to what has to be specified and tested in order to ensure that a given CMS can fulfill its purpose. Effects that are caused by physical processes within the imager like smearing and blooming rather than the camera optics are not part of the considerations. The same holds for image aberrations (defocus, distortion, image blur) and sampling artifacts (Moiré and aliasing) that will be covered elsewhere.

2.1 Origins and Occurrence

As presented in the preceding sections, the task of the optical part of an imaging system (i.e. the lens) is to represent the luminance distribution of the observed scene by a corresponding illuminance distribution on the image plane. This “sensor illuminance” is subsequently transformed into an electronic signal by the imager and finally into digital gray (resp. color-) values that can be passed on to a suitable display device (eventually via a microprocessor).

There are several reasons why the illuminance on the imager may deviate from the original luminance distribution. One prominent effect is stray light, resulting in the presence of luminous flux on areas of the imager where there is no corresponding luminance in the original scene [5]. These so-called *stray light artifacts* or ‘flares’ can have different physical reasons, like spurious reflections within the lens system, scattering or diffraction—or even a combination of all of those. It is important to note that as a matter of principle, these effects cannot be completely avoided, as for a compound lens system, some fraction of the light will always be scattered, reflected or diffracted within the objective where it shouldn’t be. The question of whether this poses a problem or not simply depends on the magnitude of the effect—more precisely, it depends on the signal to noise ratio (SNR) within the final image. Actually, two things have to be considered: First, is the spurious flare illuminance above the detection threshold of the imaging device in the first place (otherwise there will be no noise in the final signal at all):

$$E_{flare} > E_{thresh}? \tag{8}$$

Second, if so, what is the ratio between this flare value E_{flare} and the actual image illuminance E_{signal} at this point:

$$SNR \equiv \frac{E_{signal}}{E_{flare}} > SNR_{\min}? \tag{9}$$

As this kind of artifact can only *increase* the image illuminance,[2] it is obvious that this topic is especially critical for dark regions within the scene (small E_{signal}, driving at night). To get an idea of the scales involved, let us consider the case of spurious reflections.

Whenever light passes an interface between media with two different refractive indices, a fraction of it is reflected at this interface. According to Fresnel's Equations, for a glass-air-interface at normal incidence, this fraction can be calculated as:

$$R = \left(\frac{n - n'}{n + n'}\right)^2 \approx 4\% \tag{10}$$

with $n = n_{glass} \approx 1{,}5$ and $n' = n_{air} \approx 1$ (note that this value will be much higher for grazing incidence!). For a roving ray to wrongfully hit the imager, it has to be reflected an even number of times within the lens system—at least twice. Thus, for an uncoated system, this '*first-order flare*' will have a relative intensity of[3]:

$$R_{(1)} \approx R^2 \approx 0.0016 \tag{11}$$

For "higher order flares" originating from (an even number of) *multiple* internal reflections, the corresponding number is:

$$R_{(m)} \approx R_{(1)}^m \approx 0.0016^m \tag{12}$$

For a standard scene, the dynamical range of luminance values to be reproduced can easily exceed $1/R_{(1)} = 625$ (especially when using HDR cameras). At the latest in this case, i.e. if the system is supposed to be able to distinguish between more than 625 different gray values, simply using Eq. (1) and choosing the quantization steps in a way that $E_{flare} < \Delta E_{quantization}$ is not a valid option anymore. Even for smaller dynamical ranges, a spurious flare illuminance can easily become detectable if the observed light source is just bright enough (maybe not being measurable itself anymore, as it will drive the system into saturation—but still being able to cause stray light effects in an otherwise dark scene). Therefore, the use of specially coated lenses is advisable. With proper antireflective coating, the residual reflectance can easily be reduced to $R < 0.001$ for a single wavelength, resulting in $R_{(1)} < 1E\text{-}6$.

[2]Coherent effects which may actually result in a *reduction* of image illuminance by *destructive interference* are not considered.

[3]The fact that the reflected spurious ray has to pass through a larger number of lens elements than a regular ray and therefore will additionally suffer from a higher *transmission loss* has been neglected.

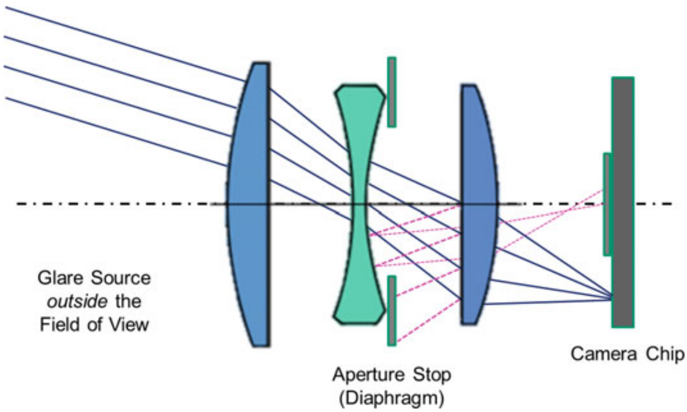


Fig. 3 Typical glare beam path showing that spurious rays may still hit the sensor even if the actual glare source itself lies outside the field of view

However, for the application in a CMS, a broadband coating for the complete visible spectrum has to be applied. For this, values of $R < 0.005$ are feasible, resulting in $R_{(1)} < 2.5E\text{-}5$.

Even with the use of AR coated lenses (and imagers resp. imager cover glasses!), there are still scenarios where flare artifacts may become visible: When a bright light source like the sun or the headlights of an approaching or trailing vehicle is shining directly into the image system (so called ‘glare scenarios’). The luminance of these ‘glare sources’ can easily be more than one million times the luminance of the actual scene to be displayed. In this context, it is important to note that it is not even necessary for the ‘glare source’ itself to be part of the visible image! Even if the image of the glare source were located outside the active sensor area, the spurious reflected rays can still hit the image plane (see Fig. 3).

As a general rule, it is always advisable to prevent light that is not part of the image formation process from entering the imaging system in the first place. Therefore, a proper lens hood in front of the camera lens can help minimize this effect. However, if the image of the glare source actually is part of the active sensor area, this does not help, as in this case the primary rays of the glare source cannot be blocked by the lens hood without obscuring the field of view.

2.2 *Types of Flares*

As indicated above, stray light artifacts may have different causes as well as different appearances in the final image. Typically, the notion *flare* or *lens flare* is used as an umbrella term for all kinds of stray light artifacts. With respect to the manifestation and the characteristics of the effect in the image, a distinction is made between the following types of flare:

- *Veiling Glare*
- *Directed Flares*
- *Aperture Ghosts*
- *Ghost Images*

The distinguishing features on which this classification is based are the spatial properties of the artifacts. Flares that are well localized within the image and manifest themselves in characteristic shapes are called *ghosts* or *ghosting*. Needless to say, effects like these can cause far more critical misinterpretations of the scene than a spurious illuminance that is more or less evenly distributed over the image. The different flare types will now be discussed in detail.

To illustrate the physical origin and distinct characteristics of the effects, typical beam paths will be displayed for each single type of flare. A standard Cooke Triplet is chosen for this purpose as a simple canonical imaging system. As a reference, Fig. 4 shows the ideal undisturbed beam path for the imaging of a distant object.

2.2.1 Veiling Glare

From an optical point of view, the task of imaging can be boiled down to the 're-uniting' of rays originating from a given object point at the corresponding image point. To achieve this, the rays have to be steered and bended by the lens system accordingly. From this point of view, whenever an imaging system is malfunctioning (as in the case of spurious reflections or internal scattering), it is quite probable that no 're-uniting' occurs at all for the affected fractions of the luminous flux. Therefore, one is likely to expect the spurious illuminance to be broadly distributed over the image plane, forming some sort of blurry 'fog' or 'haze'. This type of flare is called '*veiling glare*' (sometimes also '*veiling flare*').

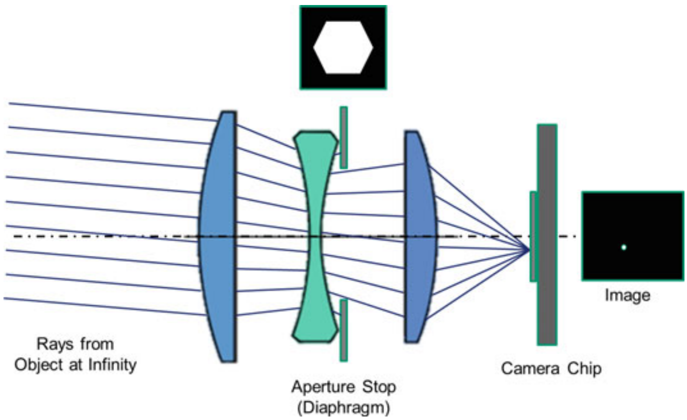


Fig. 4 Ideal beam path for the imaging of a distant object without stray light

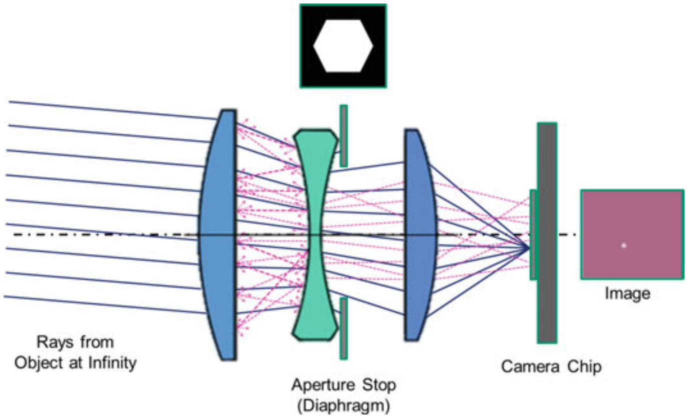


Fig. 5 Veiling glare due to uncontrolled irregular scattering within the system

As in the case of veiling glare, the imaging (focusing) capabilities of the system are supposed to fail completely for the stray light illuminance, it can be caused by any kind of spurious scattering that may be induced by surface imperfections like scratches, contamination or fogging. Figure 5 shows a corresponding beam path. However, also 'clean' residual reflections (due to insufficient AR coating) may cause veiling glare under certain conditions. In this case, one has to keep in mind that the imaging capabilities of the system do not break down completely—the additional reflecting surfaces rather define a different (now catadioptic!) imaging system (see 2.2.3 "Aperture Ghosts"). However, if this new effective imaging system is sufficiently defocused, the resulting effect cannot be distinguished from veiling glare—see Fig. 6.

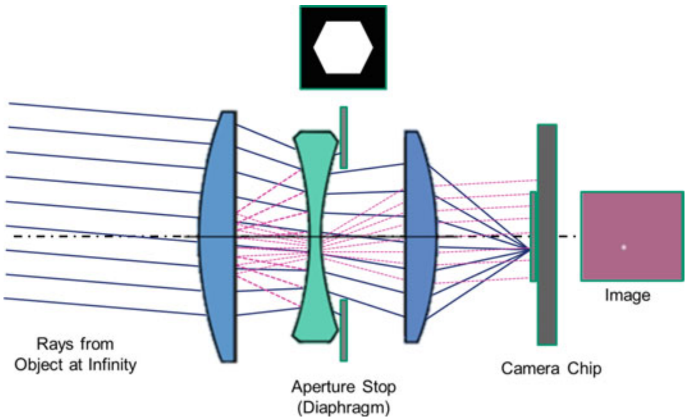


Fig. 6 Veiling glare due to spurious reflections (this corresponds to a very large diaphragm ghost covering the whole imaging sensor, see Sect. 2.2.3)

Fig. 7 Example of poor image contrast and color saturation due to veiling glare (image courtesy of Tatjana Balzer-Riesen)

The main effect of veiling glare is to reduce the image contrast. In the case of color imaging, typically also the saturation will be reduced. The reason is that the underlying scattering processes are mostly broadband effects, so that at the image point a wide range of different wavelengths is superimposed. This is also true if the underlying process is that of multiple reflections, as each image point is likely to receive stray light from different sources typically composed of different colors. Figure 7 shows an example of veiling glare.

2.2.2 Directed Flares

As opposed to the spatially extended effect described in 2.2.1, internal scattering may also result in directed effects within the image. One cause can be the localized scattering (and partly also diffraction) at the edges of the blades of the aperture stop. This results in a directed flare structure, where each single one of these flares is oriented perpendicular to the edge where the scattering resp. diffraction occurred. As variable iris diaphragms are often built from six single blades, these types of flares frequently manifest themselves as six pointed stars. The underlying beam path is shown in Fig. 8. An example of a “six pointed star” due to directed flares is shown in Fig. 9.

Note: For anamorphic lenses as they are used in cinematography, this kind of lens flare displays a strong directional anisotropy and manifests itself as a predominantly horizontal line due to the elliptical shape of the pupil.

However, directed scattering can also take place outside the aperture stop. A prominent example is the contamination of the front surface of (or any other surface within) the lens system. Quite often, these contaminations are traces of grease of smeared fingerprints. These structures act like diffraction gratings for the light passing through. Quite similar to the effect when looking at a bright light source through a smudgy windscreen that has been wiped with old blades, prominent directed flares are formed perpendicular to the grating structures. Diffraction artifacts like these can also be caused by fine scratches on the lens

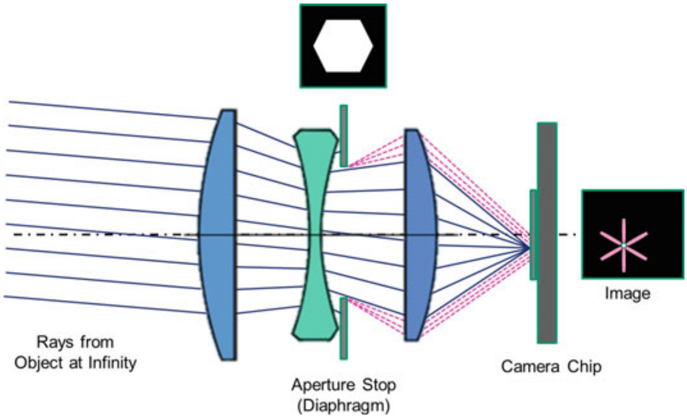


Fig. 8 Directed flares due to scattering at the aperture stop

surfaces. Manifestation and severity of those diffraction artifacts strongly depend on the coherence properties of the light involved. Therefore, when designing test setups, it is advisable to include a spatially and temporally coherent light source (pointlike/collimated and narrow band) in order to ensure that the worst case scenario is covered (Fig. 10).

At this point it should be noted that contamination of the front surface in general may pose a severe problem to CMS used in road vehicles, even if this

Fig. 9 Directed flares forming a six pointed star due to scattering at the aperture stop (image shows also some veiling glare)

Fig. 10 Directed flares due to lens contamination (image courtesy of AUDI AG)

contamination does not exhibit a periodic (grating) structure: Every small rain droplet or residual speck of road salt can create a small image of the light source somewhere that can result in a spurious high illuminance value on the chip. Therefore, the sensibility of the system against contamination should also be tested explicitly when assessing a given CMS.

2.2.3 Aperture Ghosts

As explained in Sect. 2, an important mechanism producing flares is that of multiple reflections within the lens system. If one follows the beam path for such a case, one way to look at it is to think of the system as if it had been designed that way: in this case, the light passes through an effectively catadioptric setup consisting of refractive lens elements and reflective mirror surfaces. For this special (unplanned) combination of imaging elements, there has to be an image point conjugated to the object point where the rays originally came from as well—albeit maybe not in the plane of the imager. Therefore, this indirect optical path leads to an additional defocused image on the sensor.

Now, in a defocused image, every image point is replaced by a small replica of the aperture shape—the size of this replica becoming bigger with increasing amount of defocus. If the defocus is so strong that the aperture replica is larger than the active area of the imager, we are back at the case described in Sect. 2.2.1 (*veiling glare*)—see Fig. 6. However, if the defocus of this effective imaging system is small enough, small aperture shapes may be discernible in the image (for pointlike glare sources). These shapes are called *aperture ghosts* or *diaphragm ghosts* (see Fig. 11).

For a system composed of k lenses (plus imager), there are $2k + 1$ surfaces where spurious reflections may occur. For first-order flares (see Sect. 2), there are

$$\binom{2k+1}{2} = 2k^2 + k \tag{13}$$

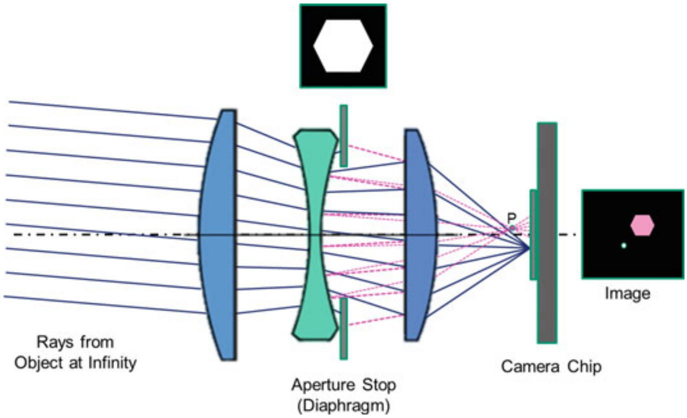


Fig. 11 Formation of a diaphragm ghost. Depending on the actual surfaces where the spurious reflections occur, the image point P is located at different positions both laterally and longitudinally. This results in different positions and sizes of the aperture ghost

ways to choose two of these surfaces (e.g. 33 for a 4-lens system). Each of these effective beam paths will result in different image locations and different amount of defocus. Where are the corresponding aperture ghosts located in the image plane? Now, one important observation is that for a rotationally symmetric system, the chief ray for a given glare source can never leave the meridional plane spanned by the glare source and the optical axis—regardless how often this ray is spuriously reflected. Therefore, the centers of all the corresponding aperture ghosts (and also of all higher order ghosts) are located on a straight line in the image plane joining the image of the glare source with the point where the optical axis pierces through the image plane. This is also true in case the actual image of the glare source lies outside the active sensor area. The underlying geometry is depicted very nicely in Fig. 2 of reference [6].

The sizes of these ghosts may vary strongly according to the different amounts of defocus involved. Figure 12 shows an example.

2.2.4 Ghost Images

As a last category of flare types, there are ghosts that actually are focused in the image plane. This results in a—typically inverted—additional image of the glare source discernible within darker parts of the original image. For this kind of artifact, almost always a reflection at the surface of the imager itself is involved. The corresponding reflected rays emerge from the image point of the glare source. Where does the second reflection have to take place in order to get a (re-)focused ghost?

Fig. 12 Aperture ghosts located on a straight line for a hexagonal diaphragm. In this example, the glare source itself (the sun) is not within the active image area. Photo courtesy of Paul van Walree [7]

There are two possibilities:

- Either the second reflection occurs at a spherical surface close to the image with the center of curvature located at the image plane.
- Or there is a second reflection at a planar surface (e.g. the rear side of a filter or a planar protection glass) in front of the lens system (if the glare source is very far away).

The first possibility is easily understood, as a spherical surface will simply re-image the back reflecting image point located at $R = 2f$ onto the chip. The second possibility is also straightforward: for a distant glare source, the rays entering the imaging system are approximately parallel. Therefore, the rays reflected back from the imager will also be parallel again after passing through the lens system in reverse. When these parallel rays are mirrored at a planar surface, they will re-enter the lens system again in a parallel way and therefore be refocused in the image plane. In digital photography, these types of ghost images are also called *filter flare* (Fig. 13).

In both cases, the ghost image will have the same size as the original image and will be inverted (i.e. upside-down) relative to the optical axis. Figure 14 shows an example.

2.3 Possible Remedies for Some Types of Flares

As most types of flares are caused by internal reflections and scattering within the imaging optics, they can be reduced by using higher quality lenses, but not completely avoided (e.g. using antireflective coating against aperture ghosts; using circular diaphragms with well fabricated blade edges against directed flares). However, if the actual glare source is located outside the field of view, the most effective counter measure would be to prevent unsolicited luminous flux not

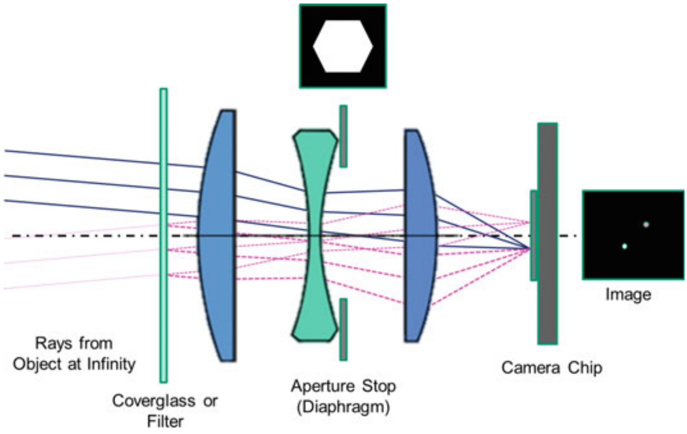


Fig. 13 Image resp. Filter Ghost caused by a planar cover glass

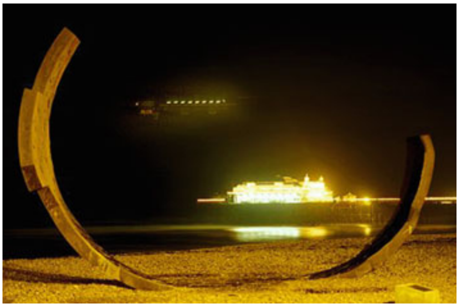


Fig. 14 Ghost image caused by a planar protective filter in front of the camera lens. Mind the point symmetry between the glare source (*lower right*) and the ghost (*upper left*). Photo courtesy of Arthur Hood and Paul van Walree [7]

contributing to the desired imaging process from entering the optical system in the first place. In some cases, this can easily be achieved by using an appropriate *lens hood*—see Figs. 15 and 16.

If a cover glass or filter has to be used in front of the imaging system, the easiest way to avoid *ghost images* is to apply a sufficient amount of *tilt* to the cover glass (if possible), see Figs. 17 and 18.

2.4 Test Scenarios

For the definition of test setups, it is important to choose the light sources used to simulate the glare sources as realistically as possible. To do so, it is not sufficient

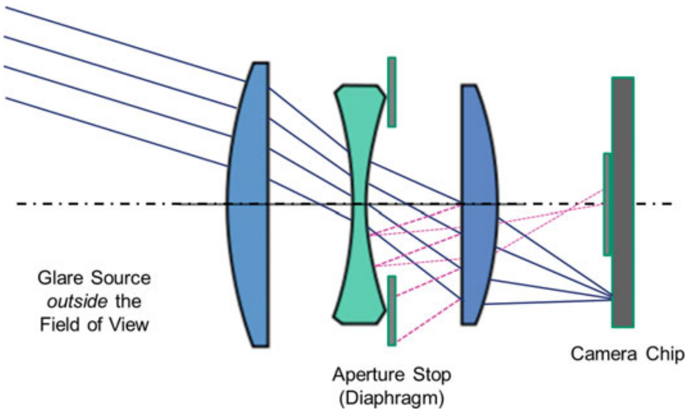


Fig. 15 Stray light artifacts by unsolicited luminous flux emerging from a glare source outside the field of view

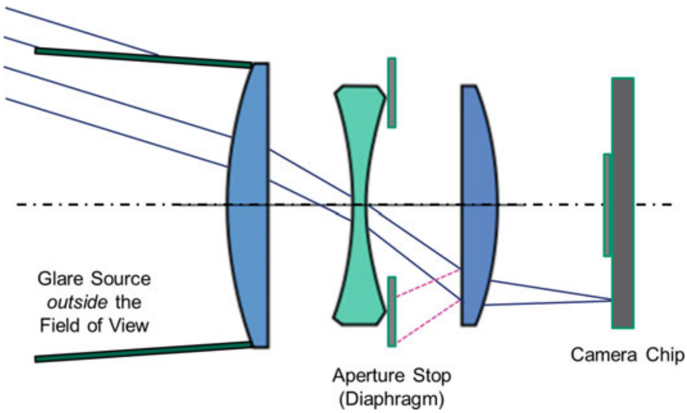


Fig. 16 Suppressing the stray light artifact by using an appropriate lens hood

just to specify a simple luminance value for the glare source. First of all, it has to be ensured that the luminous intensity actually entering the entrance pupil of the imaging system is sufficiently high. Realistically, this can only be achieved by an appropriately collimated source in order to concentrate the luminous flux into the direction towards the camera. Moreover, the direction of such a source should be adjustable. The luminous flux actually entering the optical system should be specified and monitored throughout such a test.

Secondly, different scenarios should be tested to cover all the different mechanisms described above (direct glare and indirect glare for different source locations,

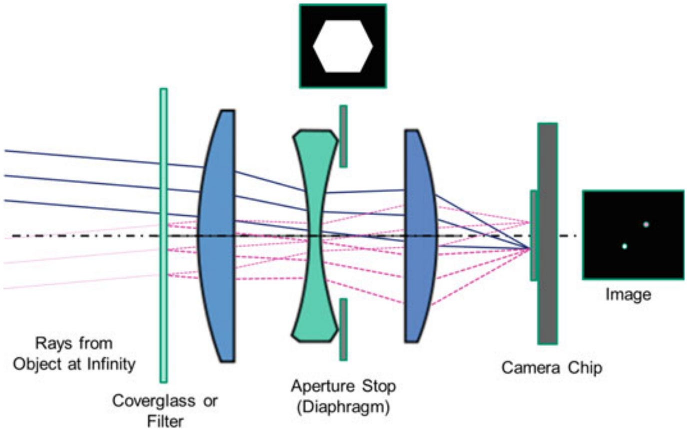


Fig. 17 Image ghost caused by a filter or cover glass

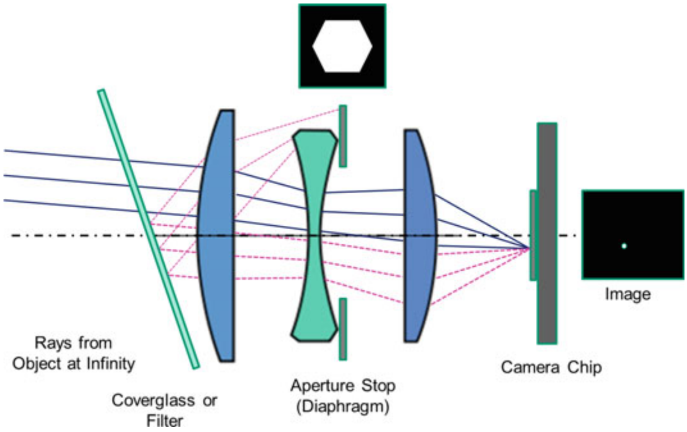


Fig. 18 Avoiding the image ghost by tilting the cover glass

different background luminances and different contamination conditions). In this context it should be checked whether the aperture settings of the system are dynamically adjusted. In this case, a lower luminous intensity for the glare may result in larger aperture ghosts, i.e. the tests should also be done for lower glare intensities.

Finally, different spectral and coherence properties of the light source should be tested, as the coatings and possible scratches/contaminations of the system may react very differently to different wavelengths.

2.5 Recap—Types of Flares and Test Setup Requirements

It's time for a short recap concerning the different stray light artifacts that can occur in glare scenarios when using camera-monitor-systems in road vehicles:

- There are four different types of flares corresponding to their appearance in the image: *veiling glare*, *directed flares*, *aperture ghosts* and *ghost images.*
- The causes of these flares can be spurious reflections, scattering and diffraction.
- As the luminance of a directly observed glare source (headlight of a vehicle shining into the imaging system) can easily be 6 orders of magnitude higher than the actual luminance of the underlying scene, flares cannot be avoided (even when using antireflective coating).
- Nevertheless, they can be minimized by using a lens hood and appropriate lens design. As a general rule, light that is not part of the imaging process itself should be prevented from entering the system.
- Test setups for the assessment of CMS should contain realistic simulation of glare sources.
- The direction of these glare sources should be adjustable and specified. As there are different mechanisms responsible for the different glare types, appropriate worst-case tests should be defined (coherent light source, direct and indirect glare etc.)
- Contamination of the front surface of the imaging system may be a major problem for CMS. Therefore, the sensitivity of the system against different kinds of contaminations (rain, mist, road salt) should be assessed when testing a CMS.

3 Propositions for Test Setups and Procedures

The goal of this section is to provide concepts and suggestions for test scenarios providing guidelines how to quantify and measure system properties regarding scattered resp. stray light (flares and ghosting). The ultimate goal is to define test procedures to ensure that a given system meets the requirements that will be defined in a corresponding ISO standard. The actual definition of corresponding threshold values for the quantities to be measured is not within the scope of this analysis.

3.1 Specifications

In Sect. 2.2, the four basic types of flares have been introduced: Veiling glare, directed flares, aperture ghosts (also called diaphragm ghosts) and ghost images (also called filter ghosts). To measure quantitatively whether and to what extent these disruptive effects are present in a given system, slightly different setups resp. evaluation methods have to be used. Before going into detail concerning the specific implementations, it is helpful to consider the common prerequisites and boundary conditions valid[4] for all the testing tasks addressed here.

3.1.1 System Accessibility and Boundary Conditions

For each test situation, the CMS under consideration has to be available within a controlled laboratory environment. This means that the system has to be removed from the car and mounted onto a specially prepared test bed. As the circumstances require, this test bed resp. certain parameters or dimensions may have to be adjusted for each specific system. For system alignment (see Sects. 3.2 and 3.4), the crucial components (camera and monitor) should be mounted in an adjustable way.

3.1.2 Interfacing

As internal system structures or signal interfaces usually won't be disclosed by the OEM, the CMS has to be regarded as an integral system ('black box'). Therefore, the available input quantity used in a test stand is limited to the luminous flux entering the entrance pupil of the camera objective. The output quantity is compulsorily given by the luminance of the display that is used to present the scene to the observer.

Accordingly, the "input section" of the test stand consists of different kinds of light sources directed at the CMS camera, creating a test scene with a specified luminance distribution. These light sources assume the roles of either the "glare source" or of a simulated "useful signal" that still has to be discernible by the system in the given test situation. In the "output section", an auxiliary diagnostic camera is used to observe (parts of) the display in a specified way. For assessment, the output signal of this diagnostic camera has to be evaluated in an appropriate way (using image processing, see Sect. 3.3). To prevent disturbances by ambient light and especially to avoid cross-talk of the light sources into the diagnostic camera, input and output section of the test stand should be optically isolated and separated from each other. Figure 19 in Sect. 3.2 shows the basic layout of such a test stand.

3.1.3 Display Properties

Unfortunately, the luminance L_{disp} of the display under test will in practice not be Lambertian. According to Sect. 1.4, this means that in general L_{disp} will be a function of both position (observed location on the display) and direction

[4] Resp. *required.*

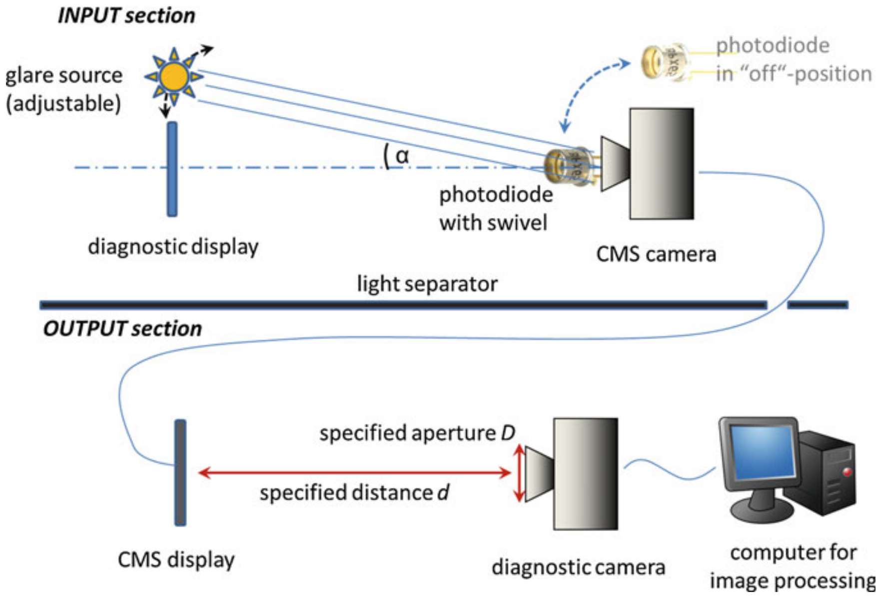


Fig. 19 Basic test setup for ghosts and flares

(observation direction): $L_{disp} = L_{disp}(x, y, \varphi, \theta)$. Especially the direction dependence normally won't be negligible. However, the pixel technology used is typically the same throughout the whole display. Therefore, position and direction dependence will at the very least decouple—in other words, the direction dependence invoked by the particular pixel technology will be the same for all pixels, i.e. positions. That means the display luminance can (to a very good approximation) be modeled by:

$$L_{disp}(x, y, \varphi, \theta) = S_{disp}(x, y) \cdot A_{disp}(\varphi, \theta) \tag{14}$$

with S_{disp} denoting the spatial (in-)homogeneity of the display and A_{disp} representing the angular (in-) homogeneity (resp. (an-)isotropy). While the variation of S_{disp} may actually be very small, this is not true for A_{disp} for state-of-the-art display technologies.

For the following considerations it will be assumed that the homogeneity of the display has already been assessed elsewhere, so that S_{disp} can be assumed constant, i.e. neglected. However, this assumption cannot that easily be made for A_{disp}, especially for low L values which are important in the current case. This would become particularly crucial if the diagnostic camera were placed very close to the monitor, as in this case different observation directions would get involved for different screen locations (although this may be calibrated in principal, but an appropriate calibration procedure would be quite cumbersome and should be avoided). This is the reason why certain guidelines concerning the diagnostic camera should be followed (presented in the next chapter).

3.1.4 Design of the Diagnostic Camera

From a theoretical point of view, the ideal choice in order to completely get rid of (i.e. *become insensitive to*) the unwanted direction dependence of the display luminance would be to utilize a telecentric lens for the diagnostic camera. However, due to the high prices of telecentric optics with large fields of view, this is not a viable solution. As an alternative, (almost) the same can be achieved by choosing the distance between display and diagnostic camera sufficiently large. In this way, the observation direction does not depend too much on the image location. This is important because for the different types of flares, signals taken from different locations on the screen have to be observed and related (see Sect. 3.3). Of course, in order to make sure that still the complete screen resp. region of interest is imaged onto the chip of the diagnostic camera (i.e. to maintain the required image ratio β), a correspondingly large focal length f has to be chosen.

For the same reasons, the aperture of the diagnostic camera should not be too large in order to avoid involving too many different directions into the imaging process.

The pragmatic approach for both the observation distance (i.e. position) and the aperture of the diagnostic camera would be to *'mimic' the human observer* (of course not concerning the image ratio β respectively the focal length f—these parameters depend on the size of the camera chip and the required region of interest!).

Following these guidelines (and using an appropriately calibrated and linearized system), the image signal of the diagnostic camera will be proportional to the display luminance L_{disp} (taken in the direction towards the observer).

3.1.5 Target Quantities

According to Sects. 1.5 and 2.1, the main effect of stray light (ghosts and flares) is to evoke unwanted display luminance at positions where there is no corresponding luminance in the original scene. Therefore, the test setups will have to measure exactly these "dark" or "background" luminances L_{black} that are due to stray light. Where to measure these values on the display will depend on the particular flare type under consideration (see Sect. 3.3).

However, after having defined what output quantity to measure and how to measure it, the question arises: *What quantity should actually be assessed?*

From an "academic" point of view, at first glance one could think of an absolute threshold value for L_{black}. However, this is not a viable way to proceed. This absolute value of L_{black} strongly depends on the current display settings which may be subject to user preferences. Therefore, the actual measure to be assessed has to be a relative one (granting the additional benefit that the proportionality factor between display luminance and diagnostic camera signal doesn't have to be calibrated exactly).

Two things have to be defined:

1. To what other luminance value should $L_{black} = L_{min}$ be related?
2. How should this relation be realized?

For both questions, there are two choices:

1. The 'reference luminance' could either be the brightest luminance L_{max} that can be provided by the display in the current setting—or it could be the particular luminance L_{useful} of the smallest useful signal that should still be discernible on the display in the given situation.
2. The relation can either be implemented by taking a simple ratio (like the common *contrast ratio CR* for displays):

$$CR = \frac{L_{\max}}{L_{\min}} \tag{15}$$

or by invoking the Michelson definition for contrast modulation:

$$C_M = \frac{L_{\max} - L_{\min}}{L_{\max} + L_{\min}} \tag{16}$$

As far as point 1 is concerned, both choices for reference luminances are important and should be assessed separately. For most practical applications, when relating L_{black} to L_{max}, the Michelson contrast is preferred, in order to avoid infinite values which otherwise would occur for an ideal system with $L_{black} = 0$. As opposed to this, when relating L_{black} to L_{useful}, there are some reasons to stick to the first contrast definition of point 2 (Eq. 15), as in this case this would exactly correspond to the well-known signal-to-noise-ratio (with $L_{black} > 0$ being the 'noise' due to flares): $SNR = L_{useful}/L_{black}$. However, it should be pointed out that the choice what definition to prefer for each case is to a certain extent arbitrary and a matter of personal taste.

To quantify only the glare effect and not the effect of an imperfect display (as it might the case that these contrast resp. SNR values are bad to begin with even in the absence of glare!), these values have to be compared for the cases with and without glare source:

$$C_M^{(noglare)} - C_M^{(glare)} < \text{thresh } (\Delta C_M) \tag{17}$$

$$SNR^{(noglare)} - SNR^{(glare)} < \text{thresh } (\Delta SNR) \tag{18}$$

In words, these criteria read as follows:

Due to glare effects, the contrast of the display signal presented to the user must not decrease more than a certain value (e.g. 0.1) and:

Due to glare effects, the signal to noise ratio for the smallest discernible signal must not decrease more than a certain value (e.g. 20)

As an alternative/extension, one could also think to introduce ‘relative’ contrast and SNR threshold values (relating the allowed threshold values to the value of the quantity without glare source), or—more or less equivalently—taking the ratio rather than the absolute difference of C_M resp. SNR for the cases with and without glare:

$$\frac{C_M^{(noglare)} - C_M^{(glare)}}{C_M^{(noglare)}} < \text{thresh} \left(\frac{\Delta C_M}{C_M} \right) \tag{19}$$

$$\frac{SNR^{(noglare)} - SNR^{(glare)}}{SNR^{(noglare)}} < \text{thresh} \left(\frac{\Delta SNR}{SNR} \right) \tag{20}$$

This is open to discussion: there are good reasons to claim that a loss of contrast of 0.1 should be considered worse if the contrast wasn’t too good to begin with.

3.1.6 Elements of the Setup

According to the points elaborated above, the test setup has to comprise the following elements:

- A light source in the input section acting as (i.e. simulating the) glare source
- A diagnostic display (preferably Lambertian) also in the input section simulating the useful signal (that should be represented on the CMS display as L_{max} resp. L_{useful} after being processed by the CMS)
- A light separator between input and output section
- A diagnostic camera in the output section observing the display
- A computer with a (simple) image processing algorithm evaluating L_{black} at the right positions (according to the type of glare) and computing the above values (see Sect. 3.3).

These elements required for each single test setup are schematically depicted in Fig. 19 (see Sect. 3.2).

3.1.7 Glare Sources

As far as the glare source is concerned, the required luminous flux cannot be provided by an isotropic point source. Instead, the light has to be collimated to a certain extent and deliberately directed into the camera system. To achieve this, the *glare source has to be aligned*.

Two glare source scenarios should be simulated:

1. The head lights of trailing car
2. The sun

For the first case, the pragmatic choice is to actually use a 'typical' car headlight. For the second case, an appropriate laser source can be used. It should be noted that in both cases it is very important to ensure that the *entrance pupil of the CMS camera is completely filled*. However, as these pupils tend to be very small, this can always (easily) be achieved by choosing a sufficiently large distance between glare source and camera.

The setup has to contain some additional diagnostic element to check whether the alignment was successful (i.e. to ensure that the active glare source observed by the system does actually have the defined luminance value in the direction towards the camera). As the glare source will have a known spatial extent, only the luminous intensity has to be checked for the given direction. This can be achieved by a photodiode that can be swiveled in front of the camera aperture. Ideally, this diode should have the same area as the aperture (resp. entrance pupil) of the CMS camera.

In Sect. 2.2.2, it was noted that a certain type of directed flares can also be caused by a contamination of the lens surface(s), acting as a diffraction grating for the incident light. Therefore, it was proposed to include a dedicated test particularly regarding the sensitivity of the CMS against contamination. As in this case the physical mechanism causing the artifact case is to a great part given by diffraction, the coherence properties of the glare source have to be taken into account. Against this background—in order to make sure that the 'worst case scenario' is covered—it is advisable to also include more coherent LED headlights as glare sources in the test.

3.2 Sample Setup

All tests can be carried out with the same basic setup depicted in Fig. 19.

The direction α along which the glare source should 'shine into the system' (and therefore the position where to mount the glare source relative to the CMS camera) depends on the particular type of glare under consideration:

- ***Veiling glare***: glare source on the axis ($\alpha = 0$)
- ***Directed flares***: glare source on the axis ($\alpha = 0$)
- ***Aperture ghosts***: large angle ($\alpha \approx 2 * \text{FoV} = 24°$), so that the image of the glare source is outside the active chip area
- ***Image ghosts***: small angle ($\alpha \approx 0.5 * \text{FoV} = 6°$)

Mind this: Aperture ghosts could to a large extent be avoided by a proper lens hood, see Sect. 2.3!

Even if it is not the primary goal of the test to distinguish between different types of flares, it is still advisable to conduct the measurements for different glare source positions to ensure that every conceivable glare scenario is covered by the test.

3.3 *Positions to Measure* $\mathbf{L_{black}}$

The positions on the screen where to measure and evaluate the parasitic background luminances also depend on the particular type of glare under consideration (see Sect. 2). They are shown in Fig. 20.

3.4 *Test Procedures*

The test procedure is divided into three different phases: preparation of the system (including the required adjustments), the actual measurements, and evaluation of the acquired data. The complete procedure can be described as follows:

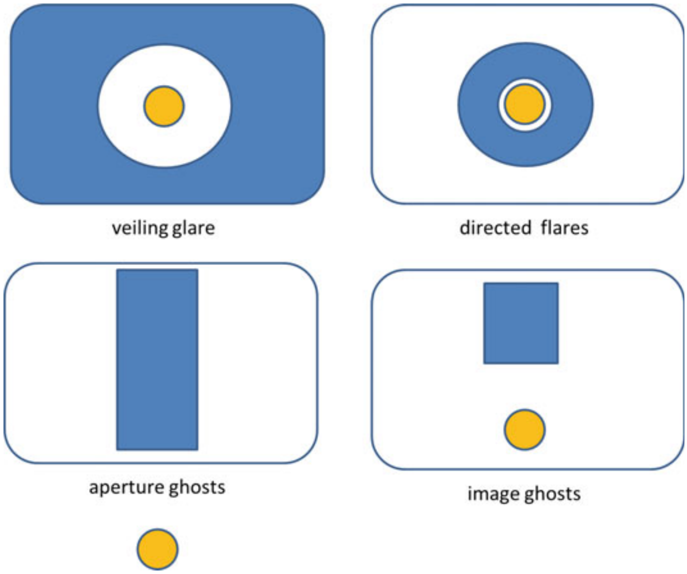


Fig. 20 Regions where to evaluate L_{black} for the different types of glare (evaluation regions depicted *blue*, image of glare sources depicted *yellow*)

1. ***System adjustment***:
   - Glare source switched on
   - Diagnostic diode in “on” position.

   Align system until specified photocurrent has been reached.
2. ***Measurement***:
   - Glare source switched off
   - diagnostic diode in “off position”
   - Measurement of $L_{black}^{(noglare)}$, $L_{useful}^{(noglare)}$, $L_{max}^{(noglare)}$
   - Glare source switched on
   - diagnostic diode remains in “off position”
   - Measurement of $L_{black}^{(glare)}$, $L_{useful}^{(glare)}$, $L_{max}^{(glare)}$
3. ***Assessment***:
   - Computation of $C_M^{(noglare)}$, $SNR^{(noglare)}$, $C_M^{(glare)}$, $SNR^{(glare)}$
   - Comparison of differences resp. ratios with given threshold values

This procedure has to be repeated for different glare scenarios, i.e. different positions of the glare source as defined in Sect. 3.2. If required, it should also be repeated for different luminance values of the glare source and the useful signal (for the case that the imaging properties of the CMS, such as the aperture D, change with input luminance) as well as different glare source types (Laser, LED-headlight—see Sect. 3.1.7).

4 Summary and Conclusion

In this chapter, a survey has been given about how the optical performance of a camera-monitor-system (CMS) that is intended to be used in an automotive environment can be characterized and assessed. The main task of such a camera-monitor-system is to provide a faithful representation of the luminance distribution in the “object space” of the considered traffic scene. Unfortunately, for this particular application, the luminance values can span a huge range, as the camera optics has to be able to resolve poorly lit scenes and may at the same time be directly exposed to sunlight or the headlights of a trailing or approaching vehicle. For this reason, stray light artifacts of the camera optics play a crucial role in defining resp. limiting the overall performance of the system.

In order to find a consistent way to assess these systems objectively, the physical origins and the specific manifestations of different types of stray light artifacts (*veiling glare*, *directed flares*, *aperture ghosts* and *ghost images*) have been analyzed. Based on these considerations, target quantities suitable as performance indicators as well as a canonical test setup and corresponding test procedures have been proposed.

In order to apply these procedures, appropriate limits and threshold values of the proposed quantities have to be defined. As the artifacts considered here can never be completely avoided, and at the same time the human visual perception system has proven to be enormously capable of adapting to new conditions, it is advised that these values should be based on physiological tests.

Acknowledgments We thank Paul van Walree for permitting the use of the photos in Figs. 12 and 14. The main part of this work has been conducted as part of a joint project of AUDI AG and the OSMIN group at the Institute of Optics, Information and Photonics, University of Erlangen-Nuremberg. We thank Prof. Dr. Gerd Häusler for interesting and helpful discussions.

References

1. DIN 5031-2:1982-03, Optical radiation physics and illuminating engineering; evaluation of radiation by different detectors
2. DIN 5031-3:1982-03, Optical radiation physics and illuminating engineering; quantities, symbols and units of illuminating engineering
3. Wikipedia, http://en.wikipedia.org/wiki/File:Luminosity.png. Accessed 11 Feb 2015
4. ISO 16505:2015, Sec. 7.8.2 Luminance and contrast rendering, Test 3: Night condition
5. Kuwabara, G.: On the flare of lenses. J. Opt. Soc. Am. **43**(1), 53–57 (1953)
6. Koreban, F., Schechner, Y.: Geometry by deflaring. In: Proceedings of the. IEEE International Conference on Computational Photography ICCP, San Francisco (2009)
7. van Walree P.: toothwalker.org; http://toothwalker.org/optics/flare.html. Accessed 11 Feb 2015

Camera-Monitor-Systems as Solution for the Automotive Market

Mark Müller

Abstract This chapter describes your opportunity to realize camera-monitor-systems (CMS) using our semiconductor products on the market. Socionext Europe GmbH provides hardware, software and tools to build systems for the automotive sector. The international standard ISO 16505 describes the minimum technical requirements for a digital camera-monitor-system to ensure that humans don't lose their overview—even with complex systems. The control, visualization and processing of information via a single platform which combines these tasks, can be designed to be both comfortable to use and safe. In order to implement a camera-monitor-system, more than just a simple camera and an industrial monitor is required. The safety aspect must not be neglected. An intelligent controller which monitors the application and guarantees the integrity of the information that is presented is necessary.

Keywords Graphic controller · Cameras · Displays · APIX® · System on chip · Ashell

1 Introduction

Drivers expect a system to assist them and be immediately available at the turn of the ignition key. The intuitive ease of use of these systems and equally, their safe use stand in the foreground for the system supplier. The application of digital systems means that humans and machines are in interaction each other—but not always successfully.

Solutions in display systems must be designed to be manageable, especially where the control of a vehicle is concerned.

M. Müller (✉)
Socionext Europe GmbH, Langen, Germany
e-mail: mark.mueller@socionext.com

A. Terzis (ed.), *Handbook of Camera Monitor Systems*,
Augmented Vision and Reality 5, DOI 10.1007/978-3-319-29611-1_7

Therefore products which cover a spectrum of functionality in varying degrees, and which are based on the same architecture may form the decisive step for successful business.

The automotive industry requires scalable systems for competitive implementations; this must be taken into account right at the start when selecting the semiconductor technology to use.

2 A Single Chip for Camera-Based Driver Assistance

With this technology, Socionext Europe introduces a position-sensitive, 360° all-round visibility solution, which represents a completely new approach to visual driver assistance systems. Using an algorithm developed by Socionext, the images supplied by four cameras are combined in such a way that a bird's eye view is created. To do this, the video images are seamlessly blended and projected onto a curved 3D surface. As a result, the vehicle and its environment can be seen from different angles. Using this special method, remote objects are represented in a way that is much less distorted. The visual impression differs significantly from conventional systems and makes it easier for the driver to detect obstacles in the surroundings of the vehicle safely.

The system also allows the viewing angle to be changed dynamically. Flowing virtual camera perspective movements for the driver, and the possibility to integrate an animated vehicle model in the all-around picture, are delightful features. On the other hand, this combination ensures that the driver maintains his orientation and can assess the proportions of obstacles in the vehicle's environment. This functionality allows the driver to safely move his vehicle in different situations. Parking operations in tight places are also facilitated, just as lane change or turns at confusing intersections. Another safety aspect of the system is the real time processing of motion pictures.

One such camera-based application improves the driver's overview and the vehicle's safety by collecting dead angles views and different views of the vehicle's direct environment.

The MB86R11 SoC (system on chip) is in production by Socionext to meet the high quality standards of the automotive industry and is designed for such applications [1]. It is available with a variety of new features. Essentially, each variant of this family combines a high-performance ARM® Cortex™ A9 processor in combination with an ARM® neon SIMD engine™, programmable graphic shader GPUs (Graphics Processor Unit) and the Socionext image processor.

The parallel video inputs of the devices realize various camera-based automotive applications in HD quality.

The SoC family supports RAM up to 1 GByte (DDR2-800 or DDR3-1066) as well as NAND, NOR and managed-NAND flash.

The 2D-Graphics-Engine is supported in parallel to the Socionext-3D graphics core, which is utilizes the OpenGL ES 2.0 shading language operations for the freely programmable vertex and fragment shaders. Whereas fragment shaders can alter the fragments and thus calculate the resulting pixel color, vertex shaders are used for geometric calculations and for dynamic changes to objects. The configurable architecture delivers optimized performance by the 3D-core and the 2D-engine delivers automatic 2D operations. By decoupling the rendering process, this flexible chip architecture enables to represent graphics with different frame rates and resolutions simultaneously.

A decisive advantage in automotive applications is the multi-layer concept with up to eight graphical layers per display controller output. The graphics are rendered only once and can be assigned to different displays. All derivatives of the MB86R1x family have three independent display controllers for a variety of screen resolutions, which support two units of this multi-layer concept. The single-layer concept is supported by one of the display outputs. Thus, a single SoC device can meet all the requirements for modern display applications. Also the high data throughput and fast frame rates at high resolution displays required are supported by the internal device architecture. Thus, a very flexible data display is possible when changing from one display scenario is to another. The standard interfaces utilized in the automotive industry and the communication protocols CAN, MediaLB (media local bus®), as well as Ethernet are also implemented in the hardware of the device family. The SoC family consists of three derivatives.

The MB86R11 and the MB86R13, which differ only in the CPU frequency and the MB86R12 [2].

This device is the successor of the MB86R11 and another member of the SoC family. Higher performance, a clock rate of 533 MHz instead of 400 MHz, an extended temperature range specifically for automotive applications and additional automotive functions including APIX® interfaces (3 transmitters and 1 receiver) by the company Inova Semiconductors GmbH are characteristics of the MB86R12.

Socionext Europe introduced a graphic SoC for graphics applications in the automotive industry—the world's first SoC with an integrated APIX® 2-interface.

APIX® is a standard-high-speed connection for video and control signals in various automotive applications.

Definitions:

- The term TX is used for a block which transmits pixel or audio data. This block is capable of sending and receiving control data.
- The term RX is used for a block which receives pixel or audio data. This block is capable of sending and receiving control data.
- Downstream is the link from TX to RX, means in the same direction as pixel or audio data.
- Upstream is the link from RX to TX. Only AShell or GPIO data is transferred in this direction.

Even the first APIX® generation included an EMC-optimized bit-serial transmission system for image data, making a transmission speed of 1 Gbit/s possible (via two shielded copper cables up to a distance of 20 m) [3]. In addition to the main channel, APIX® offers two bi-directional sideband channels and these can transmit control information up to 8 Mbit/s independently of the high-speed channel.

In addition to the "APIX® physical layer standard", Inova Semiconductors offers also a protocol for communication via the sideband channels. This is implemented in the automotive-shell (AShell), which regulates the data traffic over the sidebands.

The APIX2® link is a further development of the first APIX® generation system by the company Inova Semiconductors. The APIX2® version offers a data transfer rate of 3 Gbit/s for two independent video channels, which allows fast video-, audio- and bidirectional communication or control data transfer. Also, greater distances between host controller (unit send graphic) and the display unit can be bridged.

In addition to the APIX2® video high speed link, a sideband channel with a data transfer rate of 187.5 Mbit/s is available too in the second variant of APIX®. With this powerful video link, two totally different (in format and resolution) uncompressed video contents can be transmitted, for example a dual camera display system transmits in real time and at the same time. For a transfer of eight additional digital outputs, the APIX2®-link 3 Gbit/s still has enough transmission capacity available. A protection unit has been integrated in the MB88F33x for cases in which copyrighted video and audio data needs to be transmitted and unauthorized access must be prohibited, i.e. the encryption and decryption of the data with HDCP (high density content protection). Each APIX® transmission is protected additionally by a CRC (Cyclic Redundancy Check). The CRC logic uses modulo-2 arithmetic to calculate a checksum, so that the CRC module can detect errors in data blocks. Transmission errors are reported; it is also possible for security-relevant data to be sent automatically until the transfer is error-free many times with the implemented ARQ management (automatic repeat query).

The SoC MB86R12 has both APIX® options (1 Gbit/s and 3 Gbit/s) and these can be used depending on the application.

Conventional mirrors are very limited in their flexibility and are used primarily for the display of a single direction view. The trend in the automotive industry however is moving to modern display and driver information systems, based exclusively on digital displays because their unlimited flexibility allows the display of all types of information, previously unthinkable for drivers and passengers. The flexibility of such systems can be further enhanced if the display is coupled with multiple camera signals.

In addition to the standard functionalities, the device MB86R12 itself has four parallel, independent video inputs (video-capture unit). The video inputs support the formats DRGB666/888, ITU-R BT 601/656 and SMPTE 296 M (1280 × 720 px/60 fps, 1280 × 720 px/59.94 fps, 1280 × 720 px/50 fps). The internal bus architecture has been designed such that the four parallel video streams barely affect the graphics performance. One of these channels also enables the serial transfer of video via APIX2®. The hue, contrast, and brightness of the imported video images can be individually corrected using the image processor. In addition, intelligent operations regulate in specific sections of a camera image, brightness and contrast and thus improve the quality and visibility of details with under- or overexposure.

Adaptive de-interlacing (still image detection) and up-/downscaling are basic applications for the video-capture unit.

Inspired by this voracious variety of video input possibilities, the dimensions and resolution of displays are growing larger with each generation of vehicles.

The digital display as a whole system presents the semiconductor industry with new challenges, as they must provide and precisely match these properties, performance values and interfaces in their products.

In the cars of the future, the full digital display will be the primary interface to the driver and replace the conventional indicating instruments, perhaps even with direct views of the street. It will establish completely new applications in the vehicle that increase not only the level of comfort, but also the passengers' safety.

Each of the three implemented display controllers of the MB86R1x family can supply independent video content to one or more displays. The display resolutions per controller can be freely selected. Thus, it is also possible to connect typical embedded system resolutions like 1440 × 540 or 1600 × 600 pixels. Also displays with different resolution ratios (horizontal resolution x vertical resolution) based on a 133 MHz pixel clock are of course possible. Therefore each of these display controllers has its own register set and can thus generate an individual timing for different display resolutions. In addition to the output formats RGB and RSDS™ (Reduced Swing Differential Signaling) [4], the MB86R12 contains an APIX®-link module, which consists of a PHY (physical link handling hardware) for the three integrated APIX transmitters and the one APIX® receiver. Its main task is to connect the different interfaces of the PHY to the MB86R12 bus system and to the capture units and display controller. In addition, it provides a means of direct connection between the APIX® receiver (APIX-RX) and one APIX® transmitter (APIX-TX). The APIX-RX receives the video stream from an APIX® camera and the video content can be overlaid with graphic information from the MB86R12. The generated scene can be transmitted over the APIX-TX to a remote display.

3 Remote Graphic Processing Unit for Scalable Systems

Socionext Europe has been introducing new developments for graphical applications in display systems in the automotive sector on the market for years. The graphics engine, based on the sprite solution of the MB88F334 family [5] of graphics controllers (MB88F33x), is the latest development from the Socionext team; this is optimized for remote display systems of future applications.

The implementation of flexible system architecture was the target during the development of the second generation of this graphics controller family. Internal components and interfaces were developed with the aim to be able to realize automotive applications at a low cost.

Concerning graphics panels, functionality alone was in the foreground until now, but today the look and the ease of use must also convince. This becomes apparent when prompted to implement a combination of different display sizes. Like their two predecessors, the members of the MB88F33x-family are optimized stand-alone GPUs (graphic processing units) for automotive industrial applications. They are controlled by a host controller via an APIX2® interface, e.g. by MB86R12. The graphics information is transmitted via the APIX® link on the MB88F33x.

In addition to the APIX2® video high speed link, the side-band link is of great importance in the system, as full-duplex data can be sent.

The control data are data packets that are exchanged between the host controller and the peripherals of the MB88F334. The integrated APIX® block transmits the commands and parameters in ordered data packets to the corresponding address in the register area of the MB88F334. For users, the sideband link plays a crucial role in addition to the video link, because this is how the host controller gains access to the display unit. The display is both the central component and the input interface to the user, be it in a car or in an industrial control system.

All display elements can be operated from the host controller unit via the sideband link in the MB88F334 depending on the application. The MB88F334 can also transfer the brightness or temperature sensor data, conventional key information, the coordinates of the touch display or gesture recognition to the controller unit and take appropriate actions according to the evaluation of the data. Wiping, moving, rotating, enlarging or shrinking the view of a specific screen content using two-finger gestures: all of these movements are known to users of the consumer electronic products and are now expected in touch displays.

Even with the implementation of gesture control, the control data of the display must be transferred from an external controller to the MB88F334. After the co-ordinates have been processed internally, they can be transferred to the host processor via the Media Independent Interface (MII). The MII interface enables the use of the APIX2® link as an Ethernet bit transmission layer (PHY) and provides full duplex operation up to 100 Mbit/s. The sideband link therefore represents the link between the host processor, sensors and the display.

4 Graphic Controller with Control Functions

The display controller of the MB88F334 can receive images with a resolution up to Full HD. The differential data transfer RSDS™ ensures good signal quality on the display, especially in applications where difficult electromagnetic conditions prevail. If this requirement is not relevant, the display can also be controlled via an LVDS or TTL interface. To minimize electromagnetic emissions, even at Full HD, dual pixel mode is applied. Here the “trick” is to double the bus width to the display, therefore halving the display clock.

The generation of clock signals for the display can be handled either by the conventional synchronizing signals (H-Sync, V-Sync and Display Enable) or by addressing the row and column drivers of the timing controller (TCON). The freely programmable waveform of the generated timing control signals allows the emulation of virtually every common timing controller IC (TCON IC) used in display panels. The RGB data is transmitted as single-ended TTL signals or as low voltage differential swing signals conforming to the RSDS™ standard. This direct addressing mode is economically much more interesting, because no additional external control logic is required on the display side. For brightness control of the display backlight or the LED control of a dashboard, pulse-width modulated signals are used based on empirical experience. For these applications, the MB88F334 provides a total of 16 PWMs (Pulse Width Modulation units) and over 16 bit-wide registers that can be adjusted to create desired waveforms. This feature is very important for the selection of night and day viewing of the system. It selects always the driver the best view and quality of the shown pictures.

In addition to the classic display control and the previously mentioned adjustments, the MB8833x family can control additional peripheral elements. Keys and tell-tales are included as a part of the display application via GPIOs (General Purpose Input Output). Tell-Tales in the vehicle are indicators in the dashboard to inform the driver that the device is in operation or that a issue has existed with the system.

The Graphics Controller provides lean-designed internal memory. Sequences for the system configuration of the interfaces or even start-up screens can be saved in the integrated 32 KB ROM and 64 KB RAM. These images must be adjusted in size and color depth in the existing memory. If required by an application, additional external memory may be connected via the high-speed SPI (Serial Peripheral Interface). An application can normally use the hardware-based RLE compression method (Random Length Encoding) to make external memory superfluous. The small memory size of the MB88F334 distinguishes the device from classical frame buffer based graphics controllers; e.g. the MB86R12.

The MB88F334 is a line buffer based graphic controller in which the graphics, also referred to as "sprites", are composed line by line. The device accesses lists containing the image information in the chip. In this way, the cost of materials in the system can be greatly reduced, as neither an external frame buffer memory nor a wide memory bus interface must be implemented on the PCB (Printed Circuit Board). Sprites can be shown in various color depths:

From indirect color with 1, 2, 4, and 8 bpp (bit per pixel) to direct colors with 16 and 24bpp, from the internal or external memory area, in compressed or uncompressed formats-all combinations are possible for the application. The image with sprites composition is transferred after the blending and overlay with the video image of APIX® link to the display and displayed there. Before this happens, the picture quality will be improved further by controller hardware features. Methods such as "dithering" and gamma correction are applied to produce a perception of higher color depths and increased linear brightness. The dithering unit can increase the physical color resolution of a display from 5, 6, 7 or 8 bits per RGB channel to a virtual resolution of 10 bits. The physical resolution can be set individually for each channel. The 10-bit RGB output values can be spatially dithered to 8-bit resolution. The resolution is increased by mixing the two physical colors that are nearest to the virtual color code in a variable ratio either by time (temporal dithering) or by position (spatial dithering). An optimized algorithm for temporal dithering minimizes noise artifacts in the output image. The dither operation can be individually enabled or disabled for each pixel using the alpha input bit.

The color lookup table implements 3 look-up tables with 256 × 10 bit entries each. These can be used for different kinds of applications. For non-linear color transformations (called gamma correction), the color components R, G and B are processed independently. The input code is used as a table index and the table entry as a 10-bit output color code. The alpha value inputs are by-passed (unchanged) through the color lookup table. The second application is to use the table as an indirect color palette (lookup index). To implement reduced memory bandwidth, the index is interpreted to a 10-bit color in the palette.

All of these measures affect the level of graphics. But what is about the content of the shown information itself? A required by the automotive industry is the automotive safety integrity level (ASIL) for electrical items in the vehicle. The integrated signature unit in the MB88F334 monitors the defined display content and can detect a calculated CRC value of deviation in the application.

But the Signature Unit can do even more. It may also notice that the camera supplies pictures that are still valid to the application processor. If intelligent cameras have a 3 × 3 pixel large counter running in the non-visible field of the video image, the Signature Unit can detect whether the counter is still moving or if the camera or the application processor are out of order and providing a still image. In this case, the MB88F334 could turn off the camera image and provide the driver with no camera information rather than incorrect or outdated images. Whether it is ultimately due to a fault in the camera, the application processor, the cable or the connectors in use must then be determined in the garage. A prompt to visit a garage can be displayed by the MB88F334 using its internal flash memory autonomously and without a corresponding instruction from an application processors. In this way the system can signal that it has a fault and is not functioning properly—a minimal safety requirement for such a system.

5 Flexible System Architecture

The display panel can be used for various applications, e.g. as a control in vehicles, as a display in combination with a PLC (Programmable Logic Controller) in industrial production units or as an information source in a vehicle. A modular HMI, in which customized individual elements are only exchanged, plays an important role because the modular concept minimizes design risks and optimizes development time. The basic concept of a camera system consists of a host processor, which is responsible for the collection of all the video data and a display for the presentation of information. Partly with touch or gesture recognition. As a result of this modular concept, changes to the firmware of the display unit are no longer necessary. The display can now be configured in a simple manner to adapt to different requirements.

Due to its long experience and ever-growing know-how in embedded graphics applications, Socionext has developed from a semiconductor supplier to a system partner. The company's competence has expanded in all areas, right up to the implementation of entire applications and solutions.

It therefore makes sense for Socionext to cover not only panel aspects but also the requirements made on the control side. Socionext therefore offers a complete control unit in itself (SoC) in the form of the MB86R2x family of devices [6]. The SoC was developed in-house and is based on the ARM® Cortex™ A9 MPCore™ (Multicore Processor Core) along with the Imagination Technology's PowerVR™ SGX543 3D graphics engine.

This is the company's third-generation of high-performance graphic SoCs for automotive applications.

Featuring improved CPU and GPU performance for faster processing and sharper image rendering, the MB86R24 comes equipped with 6 Full HD capture input channels and 3 display output channels, allowing for greater flexibility in input/output control. The MB86R2x variants incorporate two 533 MHz processor cores for very high performance requirements.

With this chip, Socionext has been able to incorporate 'Approaching Object Detection' functionality, which notifies drivers of nearby people, bicycles and other objects using a 360° Wraparound View System that allows drivers to check their entire surroundings in 3D from any angle. In addition, the product also enables the development of integrated camera systems that consolidate and provide centralized control over a variety of onboard vehicle information units on several screens. Until now, the display of such information on multiple screens had to be controlled independently for each screen.

A powerful graphic controller driver and a dedicated BSP (Board Support Package) to make sure that a straightforward use of such multi-display applications is feasible. Graphic driver, optimized for Linux and T-Kernel OS (Operation Systems), are available for use. Support for further automotive OS's are under discussion and will be accessible beginning of 2016.

The MB86R24 is expected to significantly help in improving automobile safety, comfort and peace of mind, but also home and industrial applications that are becoming increasingly important.

In recent years, there has been growing concern regarding automobile safety, as can be seen by the establishment of the Kids and Transportation Safety Act in different regions of the world. At the same time, the demand for automotive systems is increasing on a daily basis. In light of these trends, Socionext developed the MB86R24, a graphics SoC that makes it possible to create a 360° Wraparound View System that notifies drivers about objects approaching the vehicle, as well as integrated HMI systems that connect people with data from inside and outside the vehicle.

The 360° Wraparound View System uses cameras facing forwards, backwards, left, and right to synthesize a 3D model of the environment and then to display the surroundings from any perspective. As systems like this, which give clear visual confirmation of a vehicle's surroundings, grow in popularity, there will be increasing expectations for additional functionality that reduce the likelihood of driver oversights while promoting safer, more confident driving.

The MB86R24 boasts roughly double the CPU performance and 5 times the GPU performance of its second-generation predecessors of the MB86R1x family, delivering sharper images and the ability to view surroundings from any perspective. The chip also features Approaching Object Detection, which notifies the driver of nearby objects approaching the vehicle. The proximity detection algorithm was developed jointly with Socionext Laboratories and is implemented as part of a 360° Wraparound View System.

This SoC can also take input from 6 cameras simultaneously, thereby enabling greater flexibility in rendering 3D images and making the technology applicable to a wider range of scenarios.

6 Integrated Systems

In recent years, the amount of information shared between drivers, vehicles and the outside world has been steadily growing. Such information includes battery information for electric vehicles, navigation information, connectivity with smart phones and the cloud and of course camera images from around the vehicle. Different kinds of information are displayed on different screens, such as central console displays, cluster displays, head-up displays or mirror replacement displays, all of which require separate display control. To provide such information in real time to drivers in an easy-to-understand manner, there is a need for technology that can collect information in a single location and centrally control how it is displayed depending on the driving scenario. Control systems can accomplish this and the MB86R24 is able to control each display to present information that suits the current driving scenario.

Moreover, the new SoC facilitates the development of modules and platforms for displays that can be incorporated into multiple models, rather than one-off developments for each car model as in the past. This, in turn, enables a significant reduction in component counts for display systems, while also making it easy to reuse products in different car models.

At the same time, Socionext also offers the software needed to build these systems. This enables the one-stop development of high-performance systems with far less work than before.

The 360° Wraparound View System with Approaching Object Detection works with the toolset software for the previous edition of the technology from the MB86R1x, and software enabling Approaching Object Detection functionality will also be made available.

In contrast to the MB86R12, the MB86R24 doesn't have an APIX interface. Therefore, to continue to transfer camera content and graphics to a remote display, a companion chip for APIX is necessary.

7 Socionext Unveils a Video and Communication Bridge Providing Interconnection Between an Application Processor and Automotive Interfaces

Carmakers are increasingly turning their attention to powerful application processors made by major chip manufacturers from the consumer electronics sector.

However, because these were not originally designed for automotive use, they are seldom equipped with the interfaces needed for this application. The Socionext MB86R91 APIX® companion chip communication and video bridge [7] offers vehicle manufacturers an automotive version to interconnect the growing number of in-car displays with consumer grade chips—while at the same time reducing costs. The MB86R91 APIX® Companion Chip enables the connection of modern high-performance application processors via various standard interfaces, such as single or dual OpenLDI Flat Panel Display Links and DRGB888. The fully integrated High Speed APIX2® transmitters, with a downlink data rate of 3 Gbps and an uplink rate of 187.5 Mbps, allow up to three high-resolution remote displays to be connected in parallel.

Typical automotive resolutions of up to 1920 × 720 pixels with 24-bit color depth per connection are supported, as is the transmission of touch information. The connections offer complete flexibility, allowing system architecture to use different resolutions. The integrated APIX2® receiver enables the connection of a video source whose input can be forwarded for processing to the application processor. A typical application in a vehicle would be a driver information system with freely programmable displays, which can be addressed simultaneously. Different combinations of passenger displays, control panels and central information systems, with touch screen control if required, are easy to implement. The savings that can be made with the MB86R91 APIX® companion chip on the transmitter side are also possible on the display side, thanks to Socionext's MB88F33x family, the APIX® remote graphics controllers. As the number of remote display units in vehicles continues to increase, automobile manufacturers are giving priority to reducing their cost-per-screen. Because the MB88F33x family provides "remote control" of the display, the MCU on the display side is no longer required. It also allows the display's CAN and TCON connections to be eliminated. The display is connected to your control unit via APIX® as an ultra-low-cost solution, including the full control mechanism.

The communication data are protected using the built-in AShell protocol, offering highly robust data transmission over the high speed link by error detection and retransmission mechanisms. The APIX2® devices offer full backwards compatibility to the APIX1® devices, by supporting the APIX1® physical layer as well as offering the downstream and upstream sideband functionality. Further, existing APIX1® implementations can also be connected to the device.

On connect of an APIX® transmitter to a receiver, the receiver automatically starts to 'scan' the incoming data stream of the transmitter, recovers and synchronizes to the clock and receives 'frame alignment' as soon as it successfully locked to the frame structure. Since this structure is based on various configuration options, it is necessary that transmitter and receiver are configured correctly to the same key frame parameters for link speed, data bandwidth, AShell, GPIO channels, and audio channel to reach frame alignment. All data transfers are protected by the integrated AShell protocol. APIX Automotive Shell (AShell) is a hardware implementation of an abstraction layer between the APIX Physical Layer (APIX® PHY) and its application. Main intention of the AShell is to cover the specific needs in the automotive sector in terms of reliability and security. An application, AShell and the APIX Physical Layer form a stacked bidirectional wire-based communication system. As part of this communication stack, AShell provides two interfaces for the exchange of messages from local to a remote island or vice versa. AShell is equipped with an appropriate interface to the embedding application of the APIX serial link. Each layer has interfaces to the layers above and below. In general, each layer communicates by the means of a protocol with its counterpart on the remote island. The AShell provides the services to the transmission and reception of application data ensuring data integrity and supply of information about transmission link status as well as simple error statistics.

In case that more than one type of outbound data (MII, AShell generic data, AShell support data or read/write resource access) is pending at the AShell2, a priority control is implemented and high-prioritized data can displace lower-prioritized data.

In case of application that cannot receive further data from AShell, a freeze signal is available. As long as freeze is active, payload transmission in both directions (inbound and outbound) stops. The AShell2 freeze doesn't cause data loss.

The Remote Handler, as a part of the APIX interface, unwraps transactions received from external host controller via the AShell to the hardware protocol of the AHB system bus and vice versa. Additionally, this module has the function of an intelligent interrupt controller. It builds interrupt messages and sends them as event messages to the external host controller. If an application requires reading some addresses from the local register space, every time a specific interrupt is received the Remote Handler autonomously collects the data to be read and sends this interrupt message right along with the data. In this way, the required communication overhead is reduced. In fact the AShell is transparent to the user or application and ensures the proper operation of the link and the transfer of all data. To use all of this extensive functions a software interface to hardware device is necessary without needing to know all transaction details of the MB88F334 hardware.

Socionext now offers the SoC Remote Framework, which enables the customer to use the Socionext automotive APIX® hardware acceleration to its full extent. Without any further software development, this next generation interface enables an immediate ramp up and fast startup. The Socionext SoC Remote Framework is meant as implementation sample to demonstrate how to remotely control specific devices over an Ethernet (E2IP) connection.

Due to the availability of the Remote Framework in source code the handling of the sample code is very easy to adapt to the specific characteristics of applications processor.

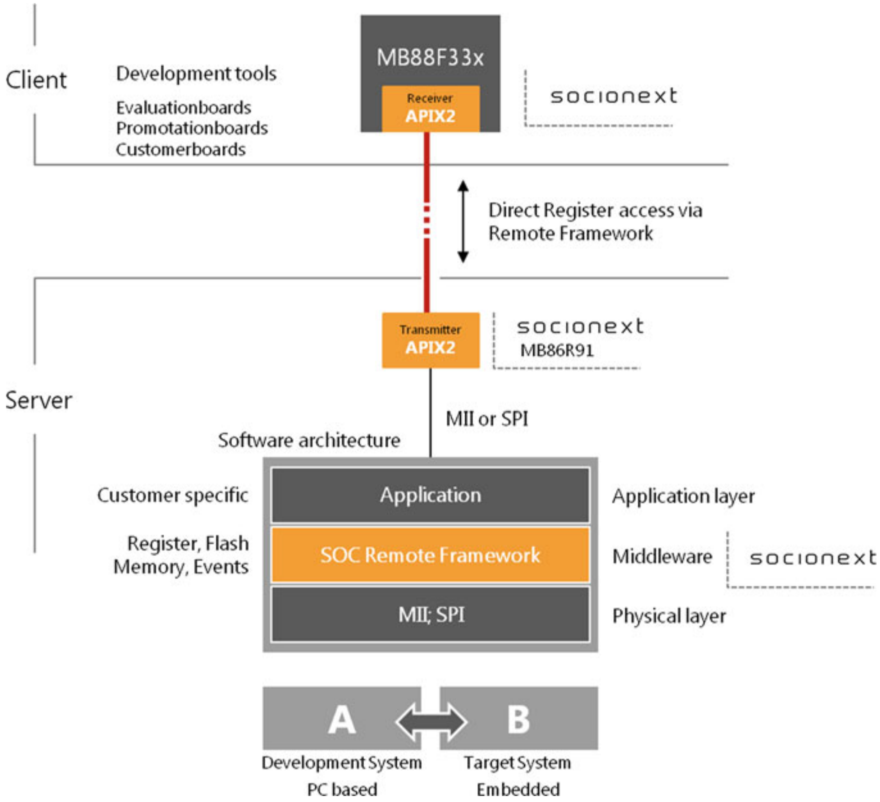


Fig. 1 Remote framework software architecture

The source code will be integrated as a part of the header- and c-file to control the accesses from the application to the MB88F33x. Both PC version for the first evaluation test and an embedded version for the serial production are available. Figure 1 shows the system structure of the SoC Remote Framework with an APIX® transmitter and the MB88F33x as APIX® remote device.

8 Multiple Applications with One APIX® Connection

The variety of today's car models is realized cost-effectively only by modular systems, with modular hardware and software components that can be reused, slightly changed, in different segments. The trigger for this change is, on the one hand, the availability of automotive-grade displays in conjunction with a rise in anticipation to the appearance and functionality.

Socionext provides the developer a compatible, scalable total solution consisting of graphic generating SoC devices and graphics receivers that are responsible for controlling the display. The portfolio ranges from entry-systems such as the MB86R12 to powerful ARM® Cortex™ A9 MPCore™ based graphics SoCs for smooth graphics display on several large TFT like the MB86R24. Camera inputs in different numbers for mirror replacement systems playing the decisive role, naturally.

Of course, the end customer will decide whether the APIX® Technology is used in the application.

In the following, two systems with APIX® Video Links for display and their flexibility will be described.

Figure 2 shows the MB86R12 SoC in a system with a camera input and a display output. Only one input and one output are shown but 4 video inputs and 3 display outputs are available. On the input side, a lot of applications are conceivable. From lane departure and sign recognition, over blind spot for park assistance, a rear view camera or mirror replacement cameras are realizable. The rear view camera has special requirements and will be connected via APIX® link to overcome the long distance of this camera to the SoC. The camera works as a remote camera and benefits from the features of the APIX® side band link. All other cameras can be connected directly to the MB86R12 video inputs. The centerpiece of this system is the MB86R12 which in addition controls the video inputs and the display outputs. After all the video information stored in the DDR memory, has been overlaid with additional data and blended, MB86R12 can output the generated graphic to a display interface. The position and the function of the display depend on the application and which camera information will be shown. If the location is far away from the MB86R12, the APIX® link to MB88F33x (on the display side) is

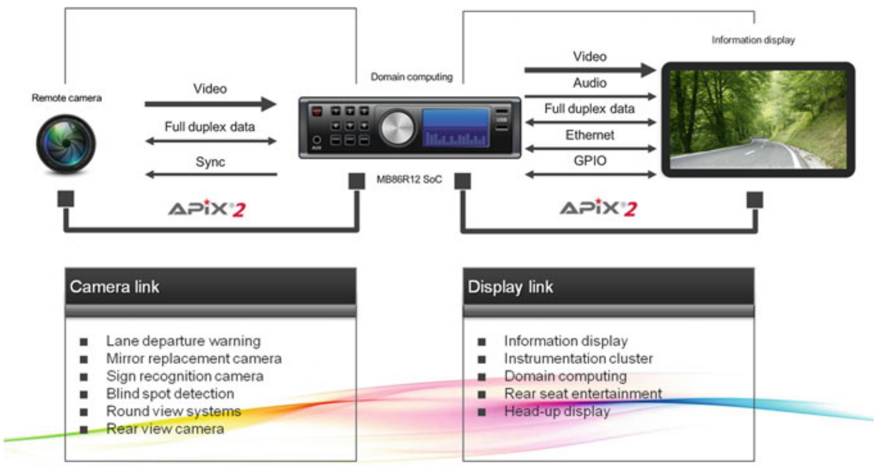


Fig. 2 3 Gbit/s fullduplex communication

necessary and will promote also the link features. In some application structures, a combination of multiple displays is feasible, e.g. it makes sense that the information from the Rear-view Camera is shown in both the information display and the head-up display. The driver can decide then which view gives him a higher level of safety.

The MB86R12 is in operation in Fig. 2. The architecture looks like more difficult but from a system structure point of view it is very simple. The MB86R91 distributes the graphic information created by the MB86R24 SoC, to the different displays in the vehicle.

The system can be split into two parts. A part which collects the camera data and generates the overlaid graphics and a part which distributes this information to different displays (Fig 3).

The MB86R24 is the heart of the functionality on the collector side. With 6 supported video capture units and the capability to connect a mobile phone, the MB86R24 is the essential unit in this arrangement. The cameras themselves could be positioned at any place in the vehicle.

For pedestrian detection, a front-view camera can detect the roadway, assisted by a radar sensor. The information is compared with stored situations. If the system detects a pedestrian, the system would decelerate accordingly depending on the

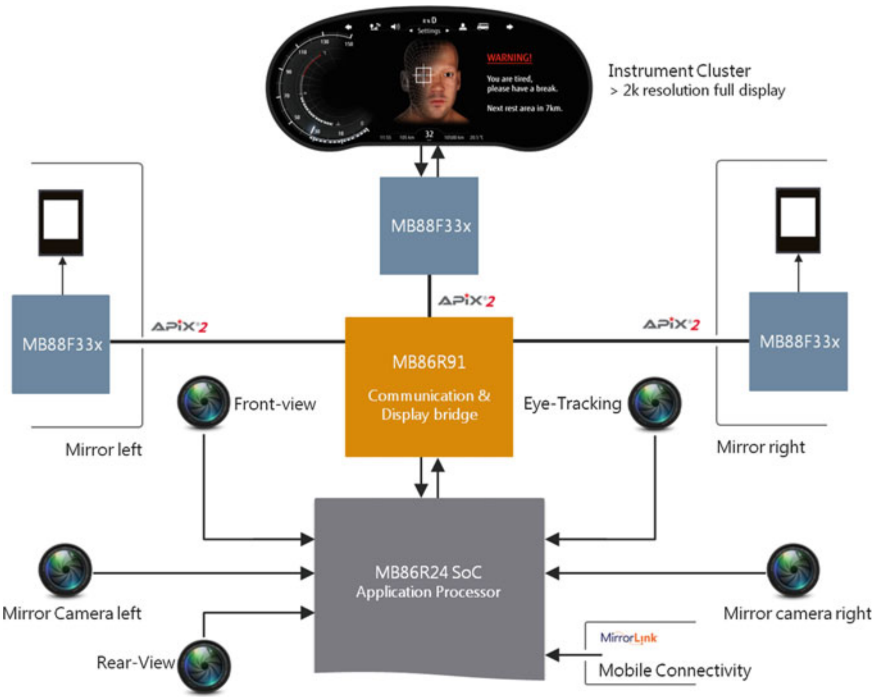


Fig. 3 All-in-one application

necessity of the situation. The rear-view camera serves to assist the driver when reversing, in particular as a parking aid. These systems have been in use for several years, especially in complex vehicles such as lorries, buses and commercial vehicles. More recently, such systems have also come into use in passenger cars.

Driving long distances on motorways involve a risk of fatigue. The motorists, truck drivers, bus drivers and travelers, often don't notice these themselves. According to studies and statistics, every fourth serious traffic accident is due to fatigue on highways [8]. This is where the attention assist function plays an important role, to which the eye tracking camera belongs. The manufacturers of attention assist follow different approaches to solve the problem. Some systems track the eyes of the driver and try to detect dangerous situations when the driver falls asleep.

And of course the mirror replacement cameras are essential in modern driver assistant systems. The possible range of that camera lens angle of more than 90 ° is just the beginning, and can sometimes even go to 170 ° upwards. Where small viewing angles that are too small prevent the avoidance of an accident, these cameras leave nothing to chance. Of course such large viewing areas incur other technical issues. The camera images with such a large horizontal resolutions show a distorted representation in a classic display panel. To eliminate this effect, but not to lose camera information, it is necessary to have a special presentation method. This is the warping method.

For this method, there is a hardware feature in MB86R24. The warping unit is a matrix inside the display controller path and the sample point can be read for each pixel from a coordinate buffer in memory. This allows any kind of static re-sampling pattern on dynamic image content, which is particularly useful for applications like lens distortion removal of wide angle cameras or windshield correction for head-up displays. By supporting this in hardware there is no impact of the warping method to the performance of the CPU or of the GPU. The image can be warped without performance loss and in real time (Warping on the fly).

The last data instructions in this system could come from a mobile phone via the MirrorLink™. This link is an interoperability standard that allows reproducing certified apps from your Smartphone on your radio or infotainment display. Software stacks on the mobile phone and the application processor enable the use of functions which improve the drivability or even existing navigation applications on the Smartphone which enhance driving comfort. After the MB86R24 has collected all the data from the cameras and other connected devices, the next job is to bring the information to the right panel. This is the job of the display controllers in the SoC. The display controllers in MB86R24 sort the graphic contents and the application decides which camera input and graphics go to which display output. OpenLDI and DRGB888 interfaces are available to transfer the data to the MB86R91, the communication and display bridge controller. With a fixed setting, the information is then forwarded to the APIX® Transmitter modules in the chip and transmitted to the MB88F33x as APIX® receiver for display.

The use of APIX® link technology makes the location of the displays freely selectable in the vehicle and provides easy use in its interior. Distances and limitations don't play a role in contrast to classic DRGB display connections. The M88F33x behind the display converts the APIX® transmitted data to the display interface format and shows the information, which the driver selects.

All the described functions allow a developer to realize an application in a short space of time, but what about the necessary PCB and devices required? Socionext Europe offers a wide range of solutions for this, starting with the evaluation board, where all the features of the chip are supported up to reference designs with optimized size and power consumption rates. This small design can be used directly in a target application and application examples are also available. But there's a lot more. Socionext offers also a service to review a customer's schematics and layouts and can help to start up his system if it is on the table.

A major challenge to use all these described features is of course the fulfillment of required standards. In the automotive environment, it is the AEC-Q100 qualification.

Automotive Electronics Council (AEC) is a standardization of the qualification of automotive electronics. The qualification for integrated circuits is summarized in the reliability test document 'Q100' and serves as a guideline for the automotive industry.

All the devices in this chapter are in-house developments of Socionext and based on solid automotive manufacturing processes and are qualified to AEC-Q100. The chips are designed for temperatures ranging from −40 to +105 degrees Celsius. One quality process for 8-D report implemented for the supply chain is understood by itself to be competitive against competitors from the consumer sector. Manufacturers of classic graphic controllers for laptops and game consoles meet these requirements only partially and especially the safety requirements cannot be met. Why should they? Because an office application or the latest strategy game attributes no significance to safety standards. Thus nobody cares if the graphics are jerky or if the operating system must be rebooted.

The trend in the automotive field determines the requirements for usable technologies. Every automobile manufacturer is currently working on the realization of autonomous driving systems.

CMS are only an intermediate step on the way to the final goal. They help the driver to keep an expanded and, from a quality point of view, always best overview. The number of built-in cameras will grow steadily, but not all images have to be displayed. Many serve the driver only as sensors for ADAS (Advanced Driver Assistance Systems) and certainly help him move in traffic. The processing of information is then no longer a human brain but is taken over by a technical brain.

For this, a controller with ever higher performance is required. This computing machine must be monitored by other controllers which provide safety functions. Thus, the spiral continues and Socionext Europe operates to develop new safety concepts for their automotive controllers.

New product variants of APIX receiver technology are in planning for the future. Also the next generation of APIX® from the company Inova Semiconductor will be the central interface in the products. In the long term Socionext will continue to develop the Graphic Controller concept, because today is clear: display resolutions above a Full HD resolution for the next generation of information displays will be a MUST.

Socionext Europe will also expand the SoC product line based on a quad-core architecture of the current ARM® Cortex™ A9 series. The goal is to offer customers the optimum solutions for a wide range of applications. In addition to other members of the High Performance Line, derivatives with quad-core CPU, products with Media Processor functionality are expected until next year. Like all Socionext products, also the graphic controllers impress even by proven high reliability and quality. The investment security for customers is ensured by the long-term availability of components and local technical support.

9 Conclusion

APIX® technology makes it possible to place the remote displays and the cameras in parts of an automobile that make it comfortable for the driver to see the necessary information to be able to drive safely. The selection of a suitable 3D Graphic processor plays a major role for processing various inputs and represents the key factor for business success. However, camera monitor systems also need a high degree of flexibility and therefore an intelligent display interface is necessary. With this approach, existing solutions can be used for implementation. Increased bandwidth requirements for ever-greater display resolutions cannot be stopped and the concept of this visual network will become more and more applicable.

APIX® is a registered trademark of Inova Semiconductors.

ARM® is a registered trademark of ARM® Corporation.

References

1. Socionext MB86R11 Hardware Manual, Rev2.40T, 14th Nov 2012
2. Socionext MB86R12 Hardware Manual, Rev,1.80, 7th Oct 2015
3. Inova Semiconductors GmbH, APIX® - 1. Generation, http://www.inova-semiconductors.de/en/technology/apix.html
4. National Semiconductor, RSDS™, "Intra-panel" Interface Specification, Revision 1.0
5. Socionext MB88F334 Hardware Manual, Rev1.40, 15th Oct 2015
6. Socionext MB86R24 Hardware Manual, Rev1.30, 26th Aug 2015
7. Socionext B86R91 Hardware Manual, Rev1.10, 19th Feb 2014
8. Wiesinger, J.: (2015), Müdigkeitserkennung durch Aufmerksamkeitsassistenten

Video Interface Technology

Rainer Gutzmer

Abstract With digitalization the number of camera interface specifications has grown enormously. The source of this growth is the technical potential of digital electronics and the extensive parameters of video technology as well as the requirements of manifold uses. Nearly every industry uses specialized camera interfaces. A short introduction to the basic parameters of video interfaces will help explain this situation. This manuscript gives an overview of camera interface technologies used in the automotive industry. The major concepts will be presented and their specific qualities described. The author comes to the conclusion that each electronic concept is purposeful for the varying requirements of a certain branch of the automotive industry. However, only the standardization of more aspects, then the electrical interface, will make cameras affordable for all thinkable automotive use cases. The author has followed the development of video technology over the course of his professional life. As a project leader at ISO, he was responsible for the automotive camera interface specified in ISO 17215.

Keywords Advanced driver assistance systems · Camera monitoring system · Camera interface · SerDes · MOST · Ethernet

List of Abbreviations

AC	Alternating Current
ADAS	Advanced Driver Assistance Systems
AUTOSAR	AUTomotive Open System ARchitecture
AHD	Analog High Definition (HDcctv Alliance)
API	Application Programming Interface
ARC	(HDMI) Audio Return Channel
ASIL	Automotive Safety Integrity Level
AVB	Audio Video Bridging

R. Gutzmer (✉)
HELLA Aglaia Mobile Vision GmbH, Berlin, Germany
e-mail: Rainer.Gutzmer@hella.com

A. Terzis (ed.), *Handbook of Camera Monitor Systems*,
Augmented Vision and Reality 5, DOI 10.1007/978-3-319-29611-1_8

BIST	Built-In Self-Test
CAN	Controller Area Network
CEC	(HDMI) Consumer Electronic Control Channel
CSI	Camera Serial Interface (MIPI)
CML	Current Mode Logic
CMS	Camera Monitoring System
CVBS	Composite Video
DDC	(HDMI) Display Data Channel
DMS	Driver Monitoring System
DVI	Digital Visual Interface
ECU	Electronic Control Unit
EHC	(MOST) External Host Controller
EMC	Electromagnetic Compatibility
EMI	Electromagnetic Interference
ESD	Electrostatic Discharge
FBlock	(MOST) Function Block
FPD	Flat Panel Display
GPU	Graphics Processor Unit
GPIO	General-Purpose Input/Output
H264	Advanced Video Coding
HDMI	High-Definition Multimedia Interface
HD	High-Definition Video
HEC	HDMI Ethernet Channel
I^2C	Inter-Integrated Circuit
IIS (I^2S)	Integrated Interchip Sound
IP	Internet Protocol
INIC	(MOST) Intelligent Network Interface Controller
LVDS	Low-Voltage Differential Signaling
MIPI	Mobile Industry Processor Interface alliance
MJPEG or M-JPEG	Motion JPEG (video compression format)
MII	Media-Independent Interface
NTSC	National Television System Committee (Analog television system)
MIPI	Mobile Industry Processor Interface
MOST	Media Oriented Systems Transport
NIC	(MOST) Network interface controller
openLDI	Open LVDS Display Interface
OSI	Open Systems Interconnection model
PHY	Physical Layer (of OSI model)
POF	Plastic Optical Fiber
RGB	Red Green Blue color space
RMII	Reduced Media-Independent Interface
SECAM	*Séquentiel couleur à mémoire* (Analog television system)
SOME/IP	Scalable Service-Oriented Middleware over IP

SPI	Serial Peripheral Interface
STP	Shielded Twisted Pair
SVS	Surround View System
TDMA	Time Division Multiple Access
TTL	Transistor-Transistor Logic
UTP	Unshielded Twisted Pair
WLAN	Wireless Local Area Network
YUV	Luminance, chrominance color space

1 Introduction

Video technology has a long history. With the appearance of ADAS and CMS, the technology is becoming more and more popular in the automotive industry.

Everyone has had experience with consumer electronics like camcorders and TVs: there is a dedicated cable; plug it in, and everything works. Someone not familiar with the details of video technology might reason that this is also the case for automotive video applications, but this is entirely untrue. The following chapter summarizes some of the fundamental reasons for the complexity of the implementation of video technology in automotive contexts. (Readers familiar with video technology might wish to skip this chapter.)

In context of this paper, a camera interface characterizes the connection of a camera with a display or computing device. This article does not cover the mechanical factors such as the mounting of the camera to the body or windshield of a car.

At first glance, an interface appears simple: a cable with one or more connectors. Unfortunately, there are uncountable approaches and practiced solutions to implementing an interface.

Which connections does a camera need to a system?

- A power supply
- A channel to deliver a sequence of images (video) and
- An interface to control the camera

Thus, a camera interface is threefold, comprising power, video and control interfaces. Since a camera interface contains a video interface, a comparison can be made to the interfaces of video player, video storage, or the monitor interface of a computer. Real-world automotive solutions will help to illustrate this aspect later in this paper.

Not so long ago there were only analog solutions for video interfaces. For power supply and control channels, separate cables were often used. Video was predominately transmitted over a 75-ohm coaxial cable. The analog standards not only determine the cable type and voltage level, but also the frame rate, number of lines, the duration of a line and the type of modulation for color information.

Worldwide, there were a limited number of standards based on CVBS: the NTSC 525-line system and the SECAM 625-line and the 625-line systems. Additionally, there were some country-specific differences in places such as France and the UK. For more details see [1]. Limited variation is the main reason why analog video interfaces were so simple to use. With the emergence of digital images, however, the number of video interface standards has exploded.

In consumer video applications, the High Definition Multimedia Interface (HDMI) has replaced analog video standards. In addition to the *display data channel* (DDC), it provides the *consumer electronic control* CEC channel, the HDMI *Ethernet channel* (HEC), the *audio return channel* (ARC) and many more. The high number of more than 10 wires in a shielded cable and the necessary royalty on each unit might be the main reason why it is rarely used in the industry.

In industrial applications, the requirements of specific applications and the commercial interests of companies combined with digital technology's capability to select nearly any aspect of a video individually has led to a confusing number of solutions. Even a look at dedicated video interfaces used for cameras shows us how very large that number is. What aspects of a video interface might be considered when selecting one of them? To answer this systematically, the following section is based on the OSI model.

2 Video Interface Basics

Video interfaces use a variety of physical layers. Most common is the use of an electrical, optical or electromagnetic medium. Since even MOST specifies an electrical interface for cameras, this paper considers only electrical interfaces.

Digital image information ranges from 8- to 24-bit data with an optional 3 bit for control signals, allowing for the use of an on-board level parallel interface. An increasing number of bits (up to 12) per pixel, higher clock rates and the demand for miniaturization have caused an increasing complexity of the PCB layout, which has led to the introduction of serial solutions at the PCB level. An Example of this is the CSI of the MIPI alliance.

For off-board cable connections, there is no alternative to serial transmission.

Serial transmission creates a demand for very high bandwidth. For instance, an uncompressed HD color video stream using YUV422 color space encoding demands a minimum of 450 Mbit/s.

$$1280 \cdot 720 \cdot 16\,\mathrm{bit/pxl} \cdot 30\,\mathrm{fps} = > 442{,}368\,\mathrm{Mbit/s} \tag{1}$$

This high amount of bandwidth requires very careful EMC considerations. Using video compression lowers the bandwidth requirements dramatically. In return the latency and quality requirements need more attention.

For example, electrical media can use either single-ended or differential signaling, and the cable needs to be chosen accordingly (Figs. 1, 2 and 3).

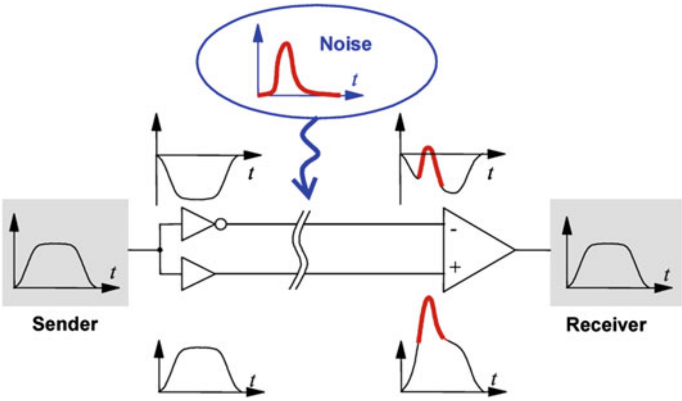


Fig. 1 Differential signaling https://commons.wikimedia.org/wiki/File:DiffSignaling.png

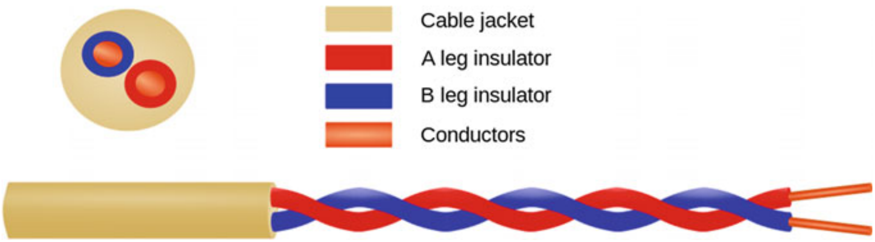


Fig. 2 Twisted-pair cable https://commons.wikimedia.org/wiki/File:Twisted_pair.svg

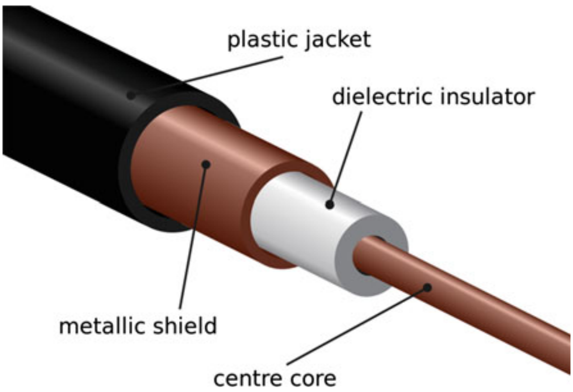


Fig. 3 Coaxial cable https://commons.wikimedia.org/wiki/File:Coaxial_cable_cutaway.svg

There has been a drawn-out discussion in the automotive industry about whether to make STP cable or coaxial cable standard. Currently, both shielded and unshielded twisted-pair cables as well as coaxial cables are in use. The latest chipsets, however, show a trend towards normal coaxial cable [2, 3].

In any video-interface standard, the type of connector is typically specified as well. For automotive applications, the type of connector is chosen by the car manufacturer.

The transmission mode can be simplex, half duplex or full duplex. While the video interface of a camera is by definition operated in simplex mode, the control interface is operated in half or full duplex.

The topology describes how the components of a system are connected with each other. According to [4], topology has both physical and logical aspects.

Nearly all nine basic topologies [5] can be found in automotive design.

Why is topology important for camera interfaces? In the early days of camera use in automotive applications, a single rearview camera and display were installed. With the introduction of SVS, CMS, DMS and many more, today platform designs can have up to ten cameras. The camera-interface technology employed limits the choice of topology, which has far-reaching consequences for system functionality and cost.

For example, automotive networks using CAN also use *bus topology* (Fig. 4). **All nodes** share the same cable, the same information, and thus the same the bandwidth. Bus topologies are widely used and cheap to implement. The necessary access control reduces available bandwidth further. But physical bus technologies are not appropriate for video interfaces.

Point-to-point topologies are the simplest to imagine. As an example, picture two endpoints, a camera and a display, connected to each other. The bandwidth is always available and exclusively used by both nodes. But physical point-to-point connections compose the backbone of modern networks with different logical topologies. Nearly all camera interfaces use this type of topology (Fig. 5).

Star topology consists of a number of point-to-point connections to a single central node. All data traffic between two nodes has to pass through the central node, so that the central node is the bottleneck of the system. When star topologies are combined in a hierarchical manner, the resulting topology is called *tree topology*. This is the way Ethernet-based networks are typically organized (Figs. 6, 7).

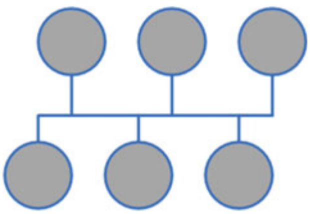


Fig. 4 Bus topology

Fig. 5 Point to point topology

In a *daisy chain*, each topology node is connected to the next in linear fashion. The traffic from node A to node C (and vice versa) has to pass through node B. This topology is used by certain camera interfaces to connect two cameras to a central node, which reduces cable use and cost (Fig. 8).

In a *ring topology*, both ends of a daisy chain are connected to each other to form a circular structure. The advantage of this topology is that if node B is down, A and C can still communicate with each other. MOST is a typical example of the use of ring topology in the automotive industry (Fig. 9).

At the application level, parameter information about the image sequence to be transmitted is required. As explained in the preceding section, for analog video this information was more or less implied by the video standard. However, as stated

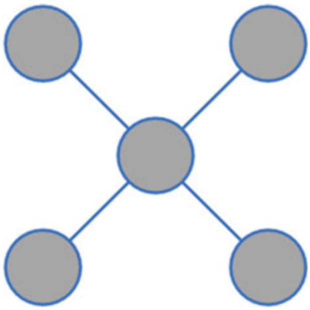


Fig. 6 Star topology

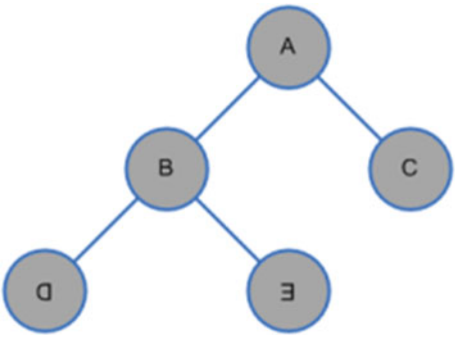


Fig. 7 Tree topology

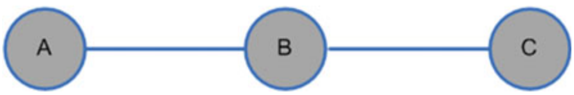


Fig. 8 Daisy chain topology

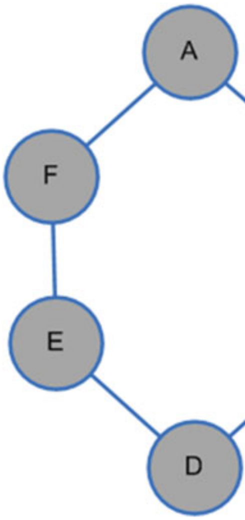


Fig. 9 Ring topology

Table 1 Typical physical media

Technical system	Medium	Connection scheme	Cable type	Number of wires	Network
Ethernet	Electrical	Symmetric	UTP	2	yes
CML solutions	Electrical	Symmetric	STP	2	no
Length reduced Gigabit Ethernet	Electrical	Symmetric	UTP	2	yes
APIX	Electrical	Symmetric	STQ	4	no
CML solutions, MOST	Electrical	Asymmetric	Coax	2	no
MOST	Optical	n/a	POF	n/a	yes
WLAN, bluetooth	Electro magnetically	n/a			yes

earlier, with the exception of broadcast video, the information is no longer provided (Table 1).

So what are these parameters? Let's start with the image. An image can be characterized by its dimensions, its height and width expresses in number of pixels. In a video, the repetition rate (frame rate) and the type of sequencing (interleaved or progressive) are other important parameters. Furthermore, the image can be monochrome or color. In both cases know of how many bits a pixel consists of. For color images, the color format (for instance, RGB or YUV) is important. Last but not least, the video stream can be transmitted either compressed or uncompressed using different methods (H264 or MJPEG) with varying settings.

All these parameters can be defined and fixed in the design. Changing conditions demand the modification of at least some of these parameters at runtime. Under these circumstances, precisely which camera parameters are controlled as well as how they are controlled must be known. Unfortunately, there are no two image sensor brands on the market that use the same method to determine this. Considering the fact that cameras from different vendors (equipped with different image sensors) are used, there needs to be a way for the operating software to recognize these differences. A well-defined camera API is able to generate an abstraction of these implementation details (Figs. 10, 11).

Such a camera API should provide information about the vendor, camera model, sensor type, calibration information and availability of feature control. The API also specifies methods for control the mode of operation, image format and certain other camera parameter. A camera API can be implemented at either end of the camera interface. This choice depends on available mounting space, component-cost considerations and cost of ownership.

However without a camera API, the integration of a particular camera brand can be anywhere from easy to impossible.

In this chapter, a device is referred to as a *camera* if the API is implemented in the camera head using a microcontroller.

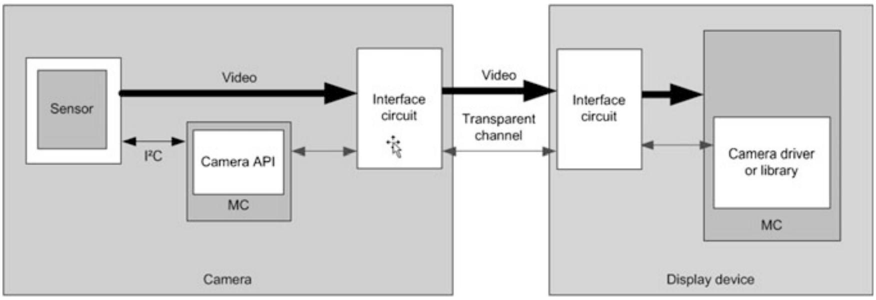


Fig. 10 Camera

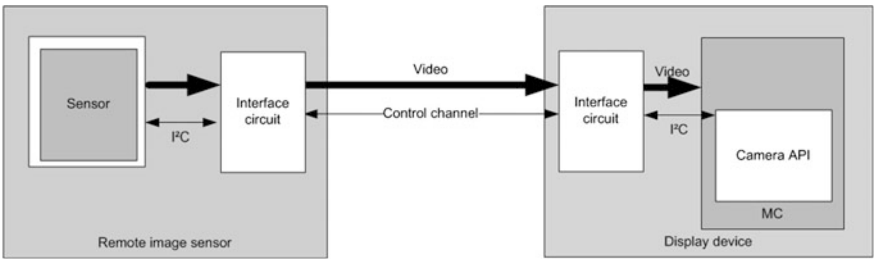


Fig. 11 Remote image sensor

A device is referred to as a *remote image sensor* if the API is implemented in the control unit (ECU).

An increasing number of applications use multiple cameras. Such applications include a CMS with more than one camera with different viewing angles or an SVS. In these applications, multiple cameras might need to capture an image at exactly the same time, which means they must be synchronized. There are two fundamental approaches to synchronizing cameras: hardware synchronization and software synchronization.

Hardware synchronization works at the image sensor. To trigger the image sensor, the clock and frame and/or line sync can be stimulated externally, or a dedicated input can be used. Although this approach is the most accurate, both image sensor and camera interface must support the distribution of the camera triggers. The minimum achievable deviation between two cameras can be measured in nanoseconds (Fig. 12).

The synchronization by software is carried out by adjusting the timing of the image sensor. For this purpose, the controlling instance calculates the expected time of the start of a certain image and compares it with the actual frame start. According to the result, the timing registers of the image sensor are modified using the camera control interface (Fig. 13).

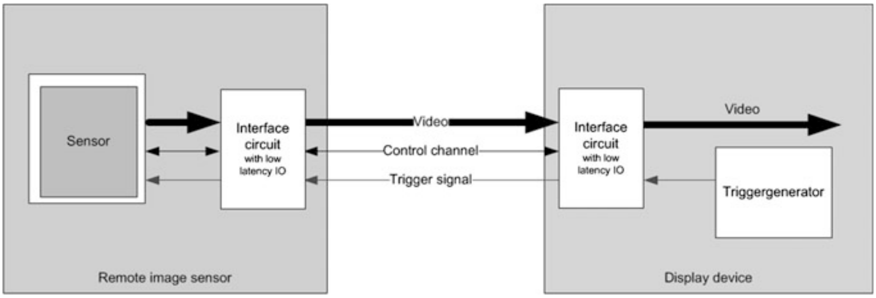


Fig. 12 Hardware synchronization

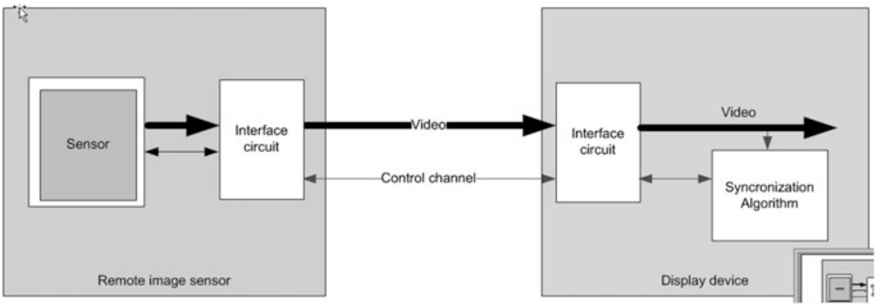


Fig. 13 Software synchronization

The minimum achievable deviation between two cameras is limited to approximately a few hundred µs. The advantage of software synchronization is that no special support from the camera interface is needed.

3 Automotive Solutions

It is important to state clearly here that, with the exception of the millions of analog rearview cameras installed in the past, there is no commonly used camera interface standard in the automotive industry today.

There are currently five camera interface solutions available on the automotive market. When implementing a CMS, one of these must be chosen. This decision is mainly influenced by the following factors:

- Development costs
- Connection costs
- Company and/or platform architecture prerequisites

Development costs include:

- Hardware development
- Software development

Camera connection costs are determined the cost of the following necessary components:

- Interface circuitry camera
- Connector camera
- Cable
- Connector display
- Interface circuitry display
- Abstraction of sensor control in the camera or the display

Additionally, considering the effect of economy of scale, the decision for a certain interface technology should take more than technical aspects into account.

3.1 Analog Cameras

Although it may seem anachronistic, analog camera solutions remain viable. Analog solutions cause negligible latency. The spectrum of analog video signals is very limited and therefore can satisfy EMC requirements easily. The length of cable that can be used in analog applications is unrivaled, and analog cameras are available from many vendors.

The resolution feasible with standard CVBS analog cameras can fulfill ISO 16505 requirements. Furthermore the analog camera interface implies the video parameter and makes the integration of a analog camera quiet easy. Interestingly, this seems to be the reason for the appearance of AHD [6] cameras in the surveillance industry. Since even displays are no longer analog the additionally effort and cost of digitizing must be calculated into connection costs.

Cable	75 ohm coaxial
Cable length	up to 300 m
Voltage level	1 Vpp
Resolution	768 · 576 aspect ratio 4:3 (1280 · 720 aspect ratio 16:9 AHD)
Frame rate	up to 30 fps

3.2 *Serializer- and Deserializer-Based Camera Interfaces*

In a point-to-point approach an image sensor is connected simply by cable to a display or image processing device. All solutions in use today were either inspired by—or directly derived from—the similarity of the video interface between ECU/GPU and a display or the video interface between a camera and an ECU. This technology is commonly used for connecting an ECU to a display even in the automotive industry.

These approaches are often referenced as LVDS solutions because these were originally serial LVDS implementations. Initially, LVDS only specified the use of a low-voltage swing on a differential signaling system. The TIA/EIA 644 does not specify bit encoding. So the data link layer differs from implementation to implementation. With the demand for higher bandwidth, all vendors of automotive-qualified silicon have changed the physical layer from LVDS to CML (Common Mode Logic). The universally well-known DVI and HDMI standards also specify a CML physical layer. Strictly speaking, these camera interfaces are based on the use of a serializer—deserializer pair. Therefore they should no longer be referred to as LVDS interfaces when they are in fact SerDes interfaces.

Common Features

The automotive industry uses SerDes implementations such as, in alphabetical order, APIX, FPD-Link and GMSL, all three of which are semiconductor, manufacturer-specific solutions. Despite their differences, they share many functionally determined features. Some of these are functionally determined, while others are determined by the requirements of the automotive industry.

Each of the three supports an embedded control channel. Since image sensors predominately use an I^2C or SPI interface for control purposes, the control channel interfaces are implemented accordingly.

Driven by the mobile-phone market, more and more image sensors and CPUs support the CSI board-level serial-video interface. So in addition to the traditional parallel video interface, this serial interface is offered for connecting the sensor to the serializer and the deserializer to the CPU.

Starting with four-wire STP cable, the use of two-wire STP cable or coaxial cable has become common. The signal interfaces to the cable are all AC coupled, which allows the transmission of power over the same cable used by the signal. In order to support the need for a variety of cable qualities and lengths, all of them have implemented some means of equalization.

Spread-spectrum technology and special line codes are used to fulfill automotive EMI requirements. The shift from advanced driver assistance to partially automated functions has forced the consideration of ASIL requirements for camera applications. Following this trend, all manufacturers have implemented features which allow the monitoring of link quality or at least the ability to do adequate diagnostics. Vendors have responded to the tendency to use SVS instead of RVC in automotive applications in the form of four-port deserializers and support for the synchronization of multiple cameras.

Automotive Pixel Link APIX™

The Automotive Pixel Link is derived from Inova Semiconductors GigaSTaR technology for long-haul display connections. The APIX physical layer is based on CML technology and features a proprietary line code with 11 % overhead. It uses an STP cable with two differential pairs. The implemented multichannel packet architecture guarantees the independent operation of several different channels. As the name implies, APIX is optimized for automotive requirements. It is chiefly used to connect an ECU—usually the head unit—to display units. For this purpose the chipsets support the transmission of many other types of information besides the video signal and necessary interfaces.

The video channel can use a parallel TTL or an open LDI interface. The audio channel uses IIS (I^2S) and the control channel uses I^2C or SPI as an interface. For low-latency purposes, GPIOs are available.

The automotive shell for secure and error-free data communication is a part of APIX specifications. The automotive shell can be implemented in both hardware and software.

In order to support IP- and Ethernet-based applications, Inova Semiconductors has added an MII/RMII interface which allows the operation of an embedded

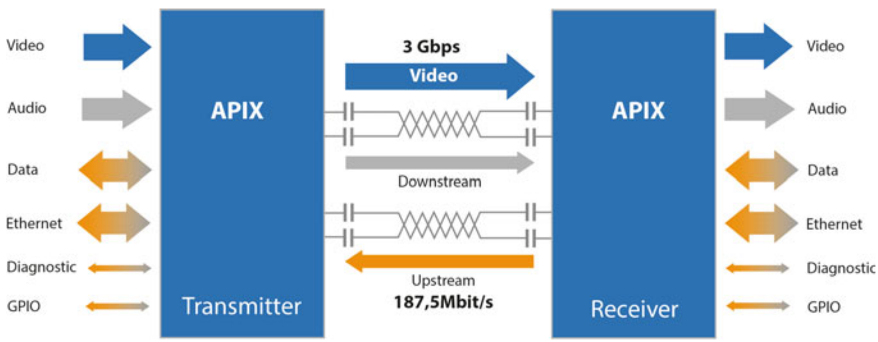


Fig. 14 APIX2 block diagram Inova Semiconductors

100Mbit Ethernet channel between ECU and display. In effect, the APIX SerDes works as a replacement of an Ethernet PHY (Fig. 14).

Summary:

Physical layer	CML
Link speed	3 Gb/s up to 12 m, 500 Mb/s up to 40 m
Video channel	10, 12, 18 and 24 Bit
Control channel	downstream embedded, upstream separate, up to 187.5 Mb/s I^2C or SPI, embedded Fast Ethernet,
Cable	Four-wire STP (two twisted pairs)
Special feature	Destination Interface parallel or open LDI, integrated MII or RMII interface to Ethernet MAC, up to 4 I^2S channels
EMI	CISPR 25 [7]

Flat Panel Display Link III FPD-Link

The FDP-Link was introduced in 1996 by National Semiconductor for the purpose of connecting the output of a computer GPU to a display-panel controller. Since then, FPD-Link has been widely used as this type of interface to internally connect a Laptop motherboard with the display. FPD-Link has been used in the automotive industry as an interface between a head unit and infotainment displays since 2001 (Fig. 15).

Second generation FPD-Link II was further optimized for automotive infotainment and camera applications. With the introduction of an embedded clock, the number of wires was reduced to only one twisted pair. Using an LVDS physical layer, the bandwidth was limited to 1 GB/s. Later FPD-Link II silicon extended this

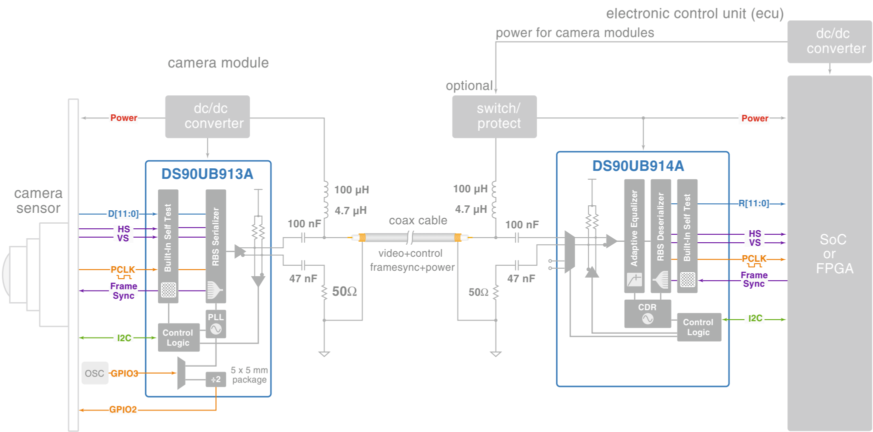


Fig. 15 FPD-Link block diagram Texas Instruments

limit by changing the physical layer to CML. With the third-generation FPD-Link III, the physical layer is exclusively based on CML, but the most significant change was the introduction of an embedded control channel.

The chipset for camera application supports error detection by providing additional bits on the line code: one bit for the forward channel and four CRC bits for the reverse channel. Errors are identified and the number of errors occurred is stored in status registers.

Summary

Physical layer	CML
Link speed	up to 3 GB/s
Video channel	12 Bit @ 75 MHz; 10 @ 100 MHz;
Control channel	embedded I^2C @ 400 kHz
Cable	two-wire STP or 50 ohm coaxial
Power consumption	250 mW

References: [2, 8, 9].

Gigabit Multimedia Serial Link GMSL (Maxim)

MAXIM introduced its serializer architecture for automotive video applications under the GMSL label in 2008. The main use case for GMSL is the delivery of video information to the displays in a car. GMSL also covers the connection of cameras to an ECU (Fig. 16).

GMSL uses CML as its physical layer. It supports both two-wire STP and coaxial cable as well as the distribution of power over the signal wires. The GMSL architecture includes an embedded control channel.

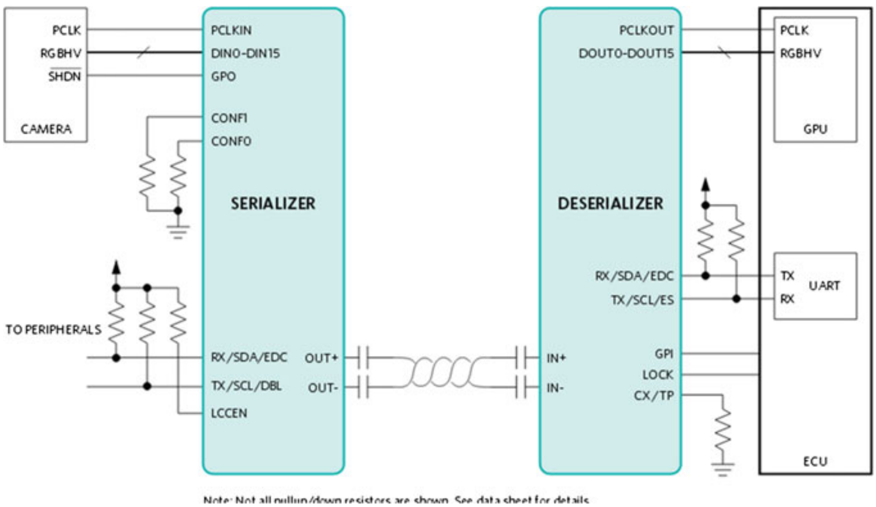


Fig. 16 GMSL block diagram Maxim

Techniques such as spread spectrum ensure compliance with the EMC requirements according to CISPR 25. The use of equalizers allows for adaption to different cable types and lengths.

The GMSL chipset for camera application supports error detection by providing additional bits on the line code. The forward channel uses one-bit parity and six-bit CRC, and the reverse channel six-bit CRC. Errors are identified and the number of errors occurred is stored in status registers.

The second generation GMSL2 is under development and will be released for production in 2016.

Summary

Physical layer	CML
Link speed	up to 3.1 GB/s up to 15 m
Video channel	8.10bit, 12bit, 16
Control channel	embedded; I^2C @ 400 kHz or UART up to 1 Mbps
Cable	two-wire STP or 50 ohm coaxial
Special feature	four-port derserializer, CSI
Power consumption	80–200 mW

Reference: [10].

3.3 Network-Based Camera Interfaces

In contrast to the SerDes-based interfaces, network-based camera interfaces were inspired by the ever-present connectivity of networks. Currently, two network-based solutions are in use and/or preparation. There are no obvious similarities in the physical layers of the two, but open specifications and second source for silicon is available.

The bandwidth is currently limited to 100–150 Mb/s. Because of the much greater bandwidth requirements of video transmission, the use of data-compression algorithms is unavoidable. With latency in mind, special considerations must be made.

The availability of low-latency compression techniques and the current development of a compatible physical layer with higher bit rates might reduce this necessity.

The complexity of a network-based camera design is generally higher than in point-to-point camera implementations.

Since the whole stack is well defined, most of the necessary software can be reused and/or bought from particular vendors.

All in all, potential market availability of compatible second-source cameras is much higher than for point-to-point interface cameras.

MOST

Media Oriented Systems Transport (MOST) is a well-established, automotive-specific technology. With a time division multiple access (TDMA) mechanism, this technology is specialized for multimedia applications. Since 2008, MOST has had an open, licensable specification which allows for second source silicon vendors. The MOST 150 specification released in 2011 adds coaxial as an alternative physical layer to the traditional plastic optical fiber (POF).

A MOST Interface NIC consists of the physical layer and low-level system services implemented in hardware. The NIC is connected to the ECU (EHC in MOST terms). The EHC executes the basic-level system services. With the increasing speed of MOST, the majority of the basic layer-system services were implemented in hardware forming an INIC (Fig. 17).

For camera solutions, a specialized INIC with integrated video interface, low-latency H264 encoder and the necessary INIC functions (including the remote control extension) are used.

The proposed camera solution consists of only two components: the image sensor and a MOST interface. The image sensor is controlled using the remote-control extension of the FBlock INIC, which is implemented as an I^2C master in the MOST interface chip.

The camera application running on the ECU accesses these I^2C masters via the MOST control channel. As in the SerDes configuration, the camera API is not implemented in the camera head. MOST recommends the use of the camera API specified in ISO 17215 for an FBlock CAMERA. For multiple camera applications like SVS, MOST offers an eight-port INIC.

Topology	ring, star for cameras
Link speed	150 Mb/s up to 100 m
Control channel	embedded; I^2C accessed over FBlock INIC
Cable	coaxial
Special feature	embedded Ethernet channel

References: [11–18].

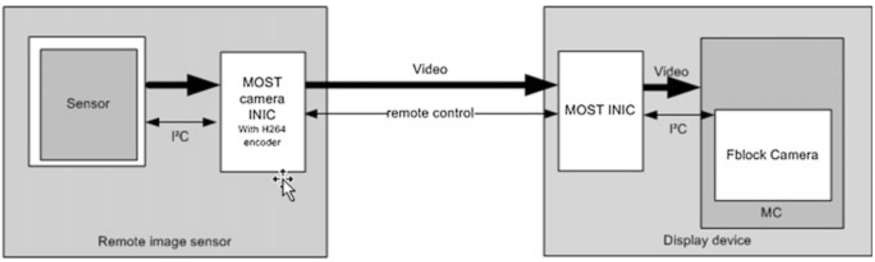


Fig. 17 MOST RIS device

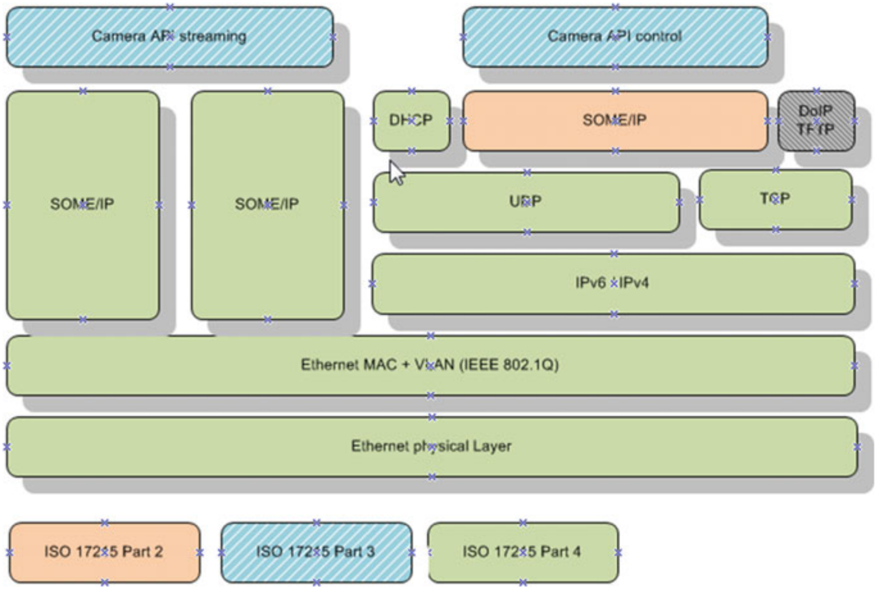


Fig. 18 Ethernet camera software stack

Ethernet
The sheer volume of current demands and development has brought Ethernet and the Internet Protocol to the attention of the automotive industry (Fig. 18).

- The time-consuming process of updating car software via traditional automotive networks has forced the development of diagnoses via Ethernet as specified in ISO 13400.
- Billing for charging an electric car will be processed over an IP connection.
- The extensively discussed car-to-X and car-to-car communication for intelligent traffic management is expected to use WLAN and the IP.
- The increasing complexity of the in-car network has initiated the idea of reducing the number of ECUs by introducing domain controller architecture. This new architecture alone creates the need for a high-performance backbone.
- The availability of a robust, cost-effective, single-wire physical interface.

In light of the events of the fourth Ethernet & IP Automotive Technology Day held in Detroit, Michigan in October of 2014, it can be concluded that Ethernet has been accepted as an automotive network of the future.

In the wake of changes, ISO 17215, *Video communication interface for cameras* (VCIC), was developed and published by ISO/TC 22/SC3/WG1. Contrary to other approaches, this standard does not specify automotive-specific technology. ISO 17215 describes how a well-proven industrial technology can be used for automotive purposes.

The precision time protocol (PTP) specified in IEEE 1588 is used to guarantee that the clock of every node in the network is perfectly synchronized. The control channel is based on SOME/IP, which is has been supported by AUTOSAR since Version 4.1. An ISO 17215 camera can be discovered as well as controlled by a SOME/IP service. For more details see [19].

Since SOME/IP and PTP are based on the IP, video information is transported via the AVB transport protocol specified in IEEE 1722. With the introduction of the BroadR-Reach® Ethernet physical layer, a connection via a UTP cable became possible. The OPEN Alliance SIG supports the process of adoption of Ethernet as an automotive network. For more details see [20]. The IEEE 802 group is working on an automotive GBit/s specification called *length reduced Gigabit Ethernet*.

A fully integrated SOC similar to the MOST camera INIC is available from Freescale Semiconductor, Inc. This SOC consist of an MC, video compression unit and on-Chip RAM, flash and an integrated Ethernet PHY. A specialized compression algorithm reduces the latency to less than one millisecond (Fig. 19).

Summary

Link speed	100 Mb/s, 1000 Mb/s under development
Video channel	streaming according IEEE1722
Control channel	part of the protocol stack
Cable	UTP
Power consumption	0.25 W

References: [2, 21–23]

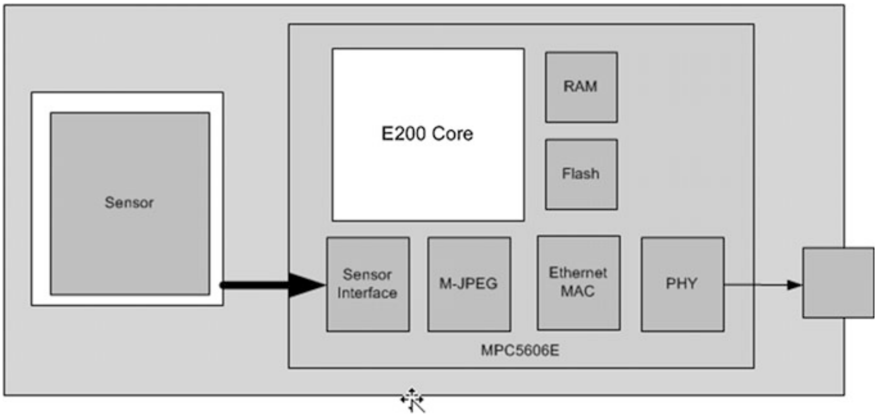


Fig. 19 Ethernet-based camera

3.4 Camera API

The ISO 17215 camera API provides specification for a set of functions and its parameters to control a camera independently of the implemented sensor. These functions can be grouped by function:

- General camera functions
- Video format functions
- Image control functions
- Imager functions

General functions give access to the so-called camera datasheet, including information about the supplier, imager, its defect pixel as well as the intrinsic and extrinsic calibration of the camera. Furthermore, the camera operation mode and startup parameters can be set and operation status read.

Video format functions are used to get information about the supported video format and region of interest, and also select one of these for current operation.

Image-control functions are intended to read and write settings, such as parameters for gain and exposure time.

Imager functions give transparent access to imager control register in the camera.

The original intention was to implement the API in the camera head. Therefore, each function is identified by its Method ID to allow for implementation over a remote connection. ISO 17215 specifies the implementation according to SOME/IP using remote procedure calls. In line with SOME/IP generic methods, functions can be grouped as follows:

- Set, get and erase functions
- Subscribe and unsubscribe functions

The first group of functions is used to modify camera parameters. The second group is used to acquire cyclical data (such as video stream) from the camera.

But the API is not limited to the use of Ethernet as a transportation medium. As mentioned previously, MOST recommends adopting this API for an FBlock CAMERA over the remote control extension of an INIC.

4 Conclusion

The positive aspect of the SerDes approach is the large amount of bandwidth available, ranging from 1 to 3 GB/s. Even when using a sensor with HD resolution, compression is unnecessary. With no compression/decompression cycle, the latency of the system is naturally very low. Serial point-to-point solutions with an embedded control channel and the implementation of ‘power over signal’ allow for the use of a two-wire cable.

Generally speaking, SerDes solutions are proprietary solutions. Even SerDes chips from the same vendor are limited in their compatibility. Considering the fact that no abstraction usually means no microcontroller is present, both camera cost and complexity of design can be reduced. In return, the ECU is not able to detect which type of image sensor is in use. On the other hand, deserializers integrated into CPUs are rare. Therefore, the display device must be equipped with a deserializer, and the software is limited to the image sensor as determined by the system design.

Should a new camera be required for repairs, the existing one can only be replaced with a camera of identical design. The average lifespan of a car in Germany is 18 years, and semiconductors are usually produced for no longer than five years. For this reason, aftermarket coverage needs to be considered carefully. Setting up a second-source supplier is also challenge.

Regardless of these facts, SerDes solutions are cost efficient in high-volume passenger car production.

Currently, MOST-based remote image sensors and Ethernet-based cameras can only use a limited amount of bandwidth, making the use of video compression inevitable. The challenge of latency is solved by the implementation of special algorithms. The resultant higher power consumption might be solved by ongoing development in semiconductor technology.

The long-term stable specification of the network-based interfaces increases the likelihood of solving second-source and aftermarket challenges.

Current attempts to specify automotive-grade Ethernet interfaces with higher transmission rates indicate a very promising future for network-based cameras.

With the given abstraction from the image sensor in use, provides throughout the camera API makes the integration of cameras much easier.

Commercial vehicles are produced in considerably lower volume than passenger cars. The cost for developing highly optimized, customer-specific cameras to this lower number of cameras will naturally lead to much higher cost per unit. The demand for replacing a camera with another camera of a different brand is more likely in this market segment. Under these circumstances, a standardized camera would be the best choice.

The agricultural machinery industry goes one step further. ISO/TC 23/SC 19 *Agricultural electronics* is working on a camera interface between tractor and implement. ISO/CD20112 focuses on the camera mechanical interface. Such an interface is also desirable for CMS cameras.

But for all camera applications in the automotive industry, a long-term, stable physical interface and an implemented standardized camera API are the preconditions to solving the challenges of second-source suppliers and aftermarket demand. With increasing volume, economy of scale will make cameras affordable in all applications and market segments.

Annex: Standards

Standard	Description
ISO 26262	Road vehicles—functional safety
ISO 10605	Road vehicles—test methods for electrical disturbances from electrostatic discharge
ISO 17215	Road vehicles—video communication interface for cameras (VCIC)
ISO 13400	Road vehicles—diagnostic communication over internet protocol (DoIP)
ISO 16505	Road vehicles—ergonomic and performance aspects of camera monitor systems—requirements and test procedures
ISO 11452	Road vehicles—component test methods for electrical disturbances from narrowband radiated electromagnetic energy
CISPR 25	Radio disturbance characteristics for the protection of receivers used on board vehicles, boats, and on devices—limits and methods of measurement
IEEE 1588	Precision Time Protocol
IEEE1722	IEEE standard for layer 2 transport protocol working group for time-sensitive streams
IEEE1733	IEEE standard for layer 3 transport protocol for time-sensitive applications in local area networks
IEC 61000 part 1-9	Electromagnetic compatibility (EMC)
ANSI/TIA/EIA-644-1995	Low voltage differential signaling standard

References

1. Jack, K.: Video Demystified, 5th edn. Newnes, Amsterdam (2007)
2. Lewis, D.: Comparing Ethernet and SerDes in ADAS Applications (2013)
3. Buntz, D., Leier, D.H., Kibler, D.T., Weller, H.: Umhüllt, nicht verdrillt, *Elektronik automotive,* vol. 2 (2012)
4. Mung, C., Yang, M.: Towards Network X-ities From a Topological Point of View: Evolvability and Scalability, 42nd Allerton Conference (2004)
5. Bisci, B.: Network Design Basics for Cabling Professionals, McGraw-Hill Professional (2002)
6. HDcctv—The Fast Path to HD Surveillance (2015). [Online] Available: http://www.highdefcctv.org/hdcctv-specification. [Accessed 20 09 2015]
7. APIX—Inova Semiconductor GmbH (2015). [Online] Available: http://www.inova-semiconductors.de/en/technology/apix.html. [Accessed 20 09 2015]
8. FPD-Link III SerDes—Products, Texas Intruments (2015). [Online] Available: http://www.ti.com/lsds/ti/interface/fpd-linkiii-ser-des-products.page. [Accessed 20 09 2015]
9. M. Sauerwald, M.: FPD-Link III– doing more with less. Anal. Appl. J. Automot (2014)
10. Gigabit Multimedia Serial Link (GMSL) SerDes ICs (2015). [Online] Available: https://www.maximintegrated.com/en/products/interface/high-speed-signaling/gmsl.html. [Accessed 20 09 2015]

11. MOST® Cooperation Releases MOST150 Electrical Physical Layer Specification (2011)
12. MOST150 Star Topology Enables Advanced Driver Assist Systems (2013)
13. MOST150 Design ermöglicht größere Streaming-Bandbreite (2014)
14. Klos, R.: MOST Quo Vadis?, MOST Forum (2013)
15. Sostawa, B.: MOST® Networking Approach for Video-Camera Systems in ADAS, MOST Forum (2014)
16. Sostawa, B.: Developing an AVM Based on an In-Vehicle Infotainment Processor, *Elektronik Automotive Special Issue MOST,* 05 (2014)
17. Bott, W.: MOST in Driver Assistance, Tokyo, Interconnectivity Conference Asia (2013)
18. Mayer, E.: MOST for transmission of multimedia data. Elektronik Automotive Serial Bus Systems in the Automobile Part **5**, 12 (2007)
19. Scalable service-Oriented Middleware over IP (SOME/IP) (2015). [Online] Available: http://some-ip.com. [Accessed 22 09 2015]
20. Open Alliance FAQ, Open Alliance FAQ, (2015). [Online] Available: http://www.opensig.org/faq/. [Accessed 22 09 2015]
21. 2014 IEEE-SA Ethernet & IP @ Automotive Technologie Day - Program (2014)
22. Basic Rear View Camera (2015) [Online] Available: http://www.freescale.com/applications/automotive/adas/basic-rear-view-camera:APLBRVCAM. [Accessed 22 09 2015]
23. 32-bit Automotive MCUs Qorivva MPC5606E fact sheet (2015). [Online]. Available: http://www.freescale.com/files/microcontrollers/doc/fact_sheet/MPC5606EFS.pdf. [Accessed 22 09 2015]

Part III
Human Visual Perception and Ergonomic Design

Human Visual Perception

Christoph Rößing

Abstract The following chapter outlines the properties of the human visual system as optical system as well as the processing stages taking place between the capturing of light in the retina and generation of a mental scene representation in the visual cortex. Therefore, the visual pathway is followed from the neuronal cascade triggered by electromagnetic stimulation of the retina, which routes through the visual preprocessor (lateral geniculate nucleus) and terminates in the visual cortex where the neuronal signal is reassembled to a mental reflection of the viewed scene. Furthermore, a closer look on the subconscious processing of depth and motion cues as well as on visual search-and-find is taken. Especially the role of lower level neuronal processing stages in the retina and the lateral geniculate nucleus and their sensitivity to pictorial cues is analyzed. Based on these findings new rendering techniques may manipulate the output of low level neuronal processing stages by utilizing pictorial cues to induce or enhance the perception of distance, velocity, or saliency.

Keywords Visual perception · Monocular depth cues · Motion cues · Visual search-and-find · Perceptual rendering

Disclaimer The following chapter is part of *Video and Image Manipulation for Enhanced Perception* from *C. Rößing* [1].

1 Introduction

The sensory organ of the HVS (Human Visual System) is the retina, a photosensitive layer on the backside of the eye. However, the formation of highly detailed mental representations of the perceived scene not only demands the capture of light information but a complex information processing-chain with multiple layers of

C. Rößing (✉)
Department of Computer Graphics, Eberhard Karls University, 72076 Tübingen, Germany
e-mail: christoph.roessing@gmx.de

A. Terzis (ed.), *Handbook of Camera Monitor Systems*,
Augmented Vision and Reality 5, DOI 10.1007/978-3-319-29611-1_9

abstraction. Therefore, the HVS not only includes the retina but another two major visual units for seeing: the visual pathway and the visual cortex.

The retina contains photoreceptors which are connected to the ganglion cells via neuronal circuits. The ganglion cells' axons form the optical nerve. Besides the optical nerve, the visual pathway encloses the optic chiasm and the LGN (Lateral Geniculate Nucleus). The main processing unit for visual perception is the terminal of the optical path: The visual cortex, which contains visual stimuli responsive cells performing the final formation of mental representations of the perceived scene (see Fig. 1). It shall be noted that the fibres of the inner (nasal) and outer (temporal) part of the retina of each eye are separated in the optic chiasm and conducted to the same terminal in the visual cortex. Thus, the left hemisphere processes the right visual field of both eyes, whereas the right hemisphere processes the left visual

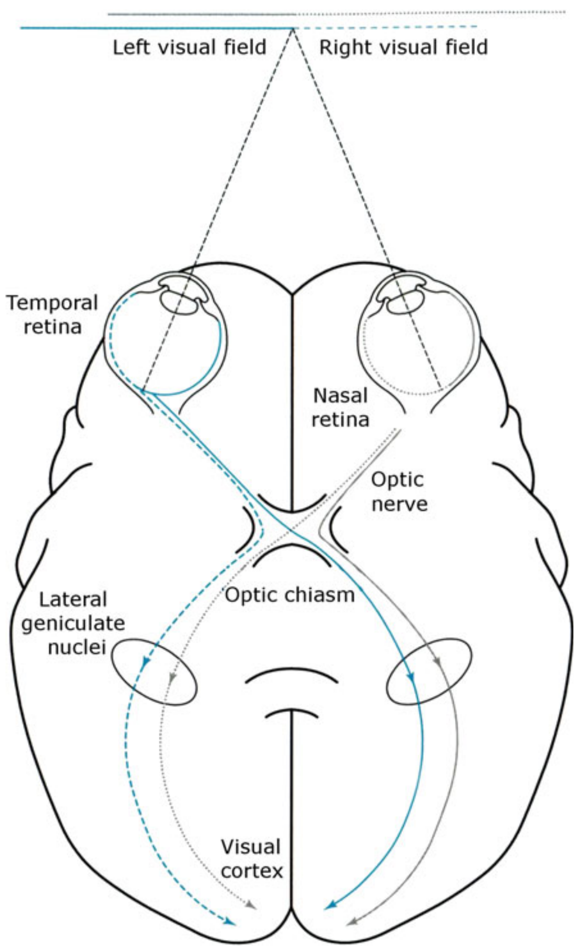


Fig. 1 The human visual system encloses the retina, the optical nerve, the optic chiasm (where *left* and *right* visual field are split up), the lateral geniculate nuclei (i.e. pre-processing stage) and the visual cortex. Reprint from [2, p. 180]

field. This enables the visual cortex to estimate distances by comparing the lateral disparity of object position in left and right visual fields (see Sect. 3.3).

2 The Human Eye

The eye with the retina is the sensory organ of vision. Its main purpose is to catch photons and direct them to the photoreceptors in the retina.

It is a light tight chamber of spherical shape which is filled with a clear viscous gel (vitreous humour). The inside surface is covered by a sheet of photoreceptors (retina). Photons pass through the cornea, the iris, and the lens to the retina. The iris is a muscular diaphragm forming an adjustable aperture followed by the lens which is refocused by the ciliary muscles (see Fig. 2).

The retina has a 1.5 mm [3, p. 13] dent in its center (fovea) in which the density of photoreceptors is increased. Therefore, this is the region with the highest spatial resolution (see Sect. 5). The optical nerve connects the retina to the higher processing stages of the visual system.

Hence, the entire human eye can be understood as an optical system with adjustable focal length (16.5–125 mm) and aperture (2–8 mm) [4]. The mean values for standard daylight vision are an aperture of ≈4, 6 mm and a focal length of 24 mm [5]. These optical values may be used to emulate the properties of the human eye for artificial depth-of-field renderings (see Chap. 15, Sect. 2).

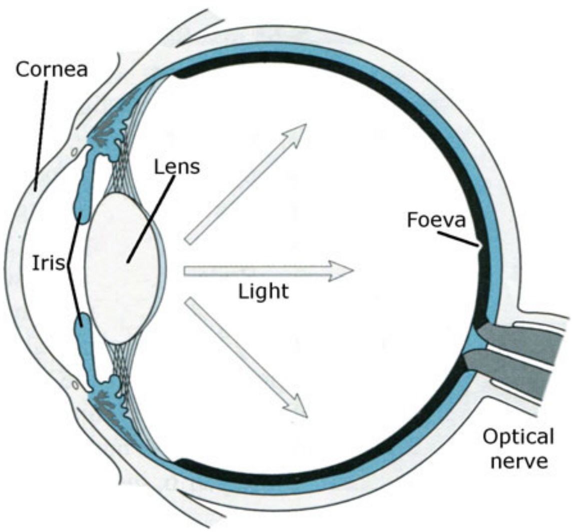


Fig. 2 The structure of the human eye. Light passes the cornea, the contractible iris, and the malleable lens up to the retina. The fovea is a small dent in the retina ($\emptyset \approx 1.5\,\text{mm}$) where the spatial optical resolution is at its maximum. Reprint of image from [2, p. 181]

3 Human Visual System

In the following, the processes taking place between initial electromagnetic stimulation in the retina and the construction of a mental representation of the surrounding scene in the cortex are discussed.

3.1 The Retina

The retina encloses three layers of different cell types: The first layer contains photoreceptors which are sensitive to light. The second inner nuclear layer interconnects these receptors with the last layer containing the ganglion cells. The latter are optimized to integrate the generated signals and to amplify neuronal responses to light impulses.

Photoreceptors
The first layer contains the photoreceptors, primary sensory cells which convert light energy into neural signals. The outer segment of the photoreceptor encloses discs filled with light sensitive pigment molecules. The inner segment contains the synaptic terminals and the cell nucleus (see Fig. 3). The terminals form synaptic clefts, which interconnect adjacent receptors via horizontal cells and ganglion via bipolar cells (see Fig. 5).

The photopigment inside the discs is rhodopsin, a protein sensitive to light. When rhodopsin is hit by photons, a chemical process named (photoisomerization) is triggered: The pigment photo-bleaches within milliseconds (see Sect. 4) and causes a change in the electrical potential of the receptor membrane. After complete

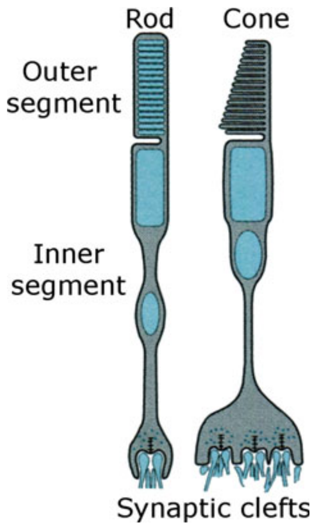


Fig. 3 Rod (high sensitivity) and cone (trichromatic color sensitivity) photoreceptors. The discs in the outer segment are filled with photopigment. The inner segment is formed by synaptic clefts which are the neuronal interface of the receptor. Reprint from [2, p. 182]

photo-induced decay of the photopigment in the disc, it is shed and absorbed into the epithelium at the tip of the receptor. Inside the epithelium, the reaction product of rhodopsin is reversed which takes up to 45 min [6]. After regeneration, the pigment is transported back into the photo sensitive segment [7]. It shall be noted that the outer cell layer of the retina is not the photosensitive segment but the synaptic terminals. Thus, light has to penetrate all other cell layers before it hits the photopigment.

Wavelength and Sensitivity

Rhodospin is a group of proteins which have their peak sensitivity at different wavelengths of light. Each cone in the human retina mainly contains only one kind of rhodopsin which makes them sensitive to specific wavelengths: S-cones have their peak efficiency at 420 nm (blue), M-cones at 534 nm (green), and L-cones at 564 nm (red). The HVS comprises 1.6 more red cones than green ones and only a small amount of blue receptors [8]. The high number of L-cones results in an increased neuronal activation in case of stimulation with red light in comparison with other colors. Cones require about 30–100 photon hits [2, p. 182] to generate a reliable change of electrical potential. However, they are capable of converting up to 50.000 photons per second [2, p. 182] into neuronal signals.

A fourth photoreceptor type are rods which contain photopigments with a peak sensitivity at 498 nm (cyan). Rods are much more sensitive to light and induce a change in neural potential up from a single photon hit. However, their neuronal output already saturates at about 100 photon hits per second.

It shall be noted that neither rods nor cones are able to see colors. They all have a significant peak efficiency at a specific wavelength (e.g. A red light source with about twice the lightness of a blue light source generates the same neuronal output in an S-cone) (see Fig. 4). Thus, the visual system interpolates the colors in subsequent processing stages by analyzing the color-opponency [2, p. 191]: The differential signal of the cone responses in respect to their spatial distribution allow the HVS to reconstruct the color representation of the observed scenery.

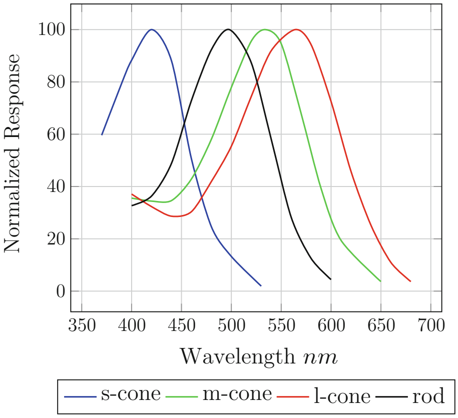


Fig. 4 Normalized responses of rods and cones at different wavelengths. (Data from Stockman and Sharpe [9])

The very light sensitive rods are excluded from color interpolation: Their neuronal signals are overlaid as monochromatic (not as color opponency) into the information feed to the optical nerve (see Sect. 3.1). From this it can be deducted why human vision fades from photopic (day light) over mesopic (twilight) to scotopic vision (low light) with a gracefully degradation from color to monochromatic vision and an increase in sensitivity [2, p. 216].

Spacing and Resolution

Most cones are located in the fovea centralis where they have their smallest size (about $1 - 4\,\mu$m in the fovea versus $\approx 4 - 10\,\mu$m outside) and spacing (0.6 arcmin). Therefore, the cone density in the fovea defines the maximum resolution of the HVS.

There are no rods located in the fovea centralis, thus light sensitivity is low compared to the rest of the receptive field. Outside the fovea rods and cones are evenly distributed but with a much higher proportion of rods ($\approx$120 million) compared to cones ($\approx$6 million). Although the number of rods exceeds the number of cones by a factor of 20, there are about 10 times more neurons devoted to the information generated by cones, which indicates that the majority of the visual information is gathered by cones inside the central field of vision.

The neuronal signals generated by the photoreceptors are the output of the first processing stage of the HVS. The number of these signals is too high to be transported via the optical nerve. In consequence, the HVS has developed strategies to preprocess and compress the gathered information in the retina.

The Inner Nuclear Layer

The smaller amount of neurons driven by the more numerous cones points to a complex wiring of these receptors located in the second layer of the retina. The inner nuclear layer connects the photoreceptors with the ganglion cells and consists of bipolar, amacrinal, and horizontal cells. Bipolar cells can be divided into ON and OFF bipolars. The OFF bipolar cells inhibit the release of neurotransmitters after a photon triggered increase in electrical potential whereas the ON bipolar cells stimulate the release of neurotransmitters. This allows the visual system to build differential signals already in the first processing stage (see Fig. 5).

Horizontal cells directly interconnect the photoreceptors and inhibit the transduction of photon energy to electrical potential in the surrounding receptors. Therefore, the transitions from dark to bright at edges or highlights cause a peak at a single photoreceptor, whereas the surrounding is dampened. This can be seen as a hard-wired high-pass filter to increase the ability of detecting high frequencies (edges, highlights, etc.).

Amacrine cells interconnect the bipolar cells horizontally with each other and vertically with the ganglion cells. Their purpose is still ill defined [10] but it is assumed that they superimpose the rod on the cone signals [11].

There are about 126 million photo receptors but only about 1 million ganglion cells [2, p. 190]. Thus, the first stage of image analysis and compression of visual stimuli is performed in the inner nuclear layer where photoreceptors connect with the ganglion cells.

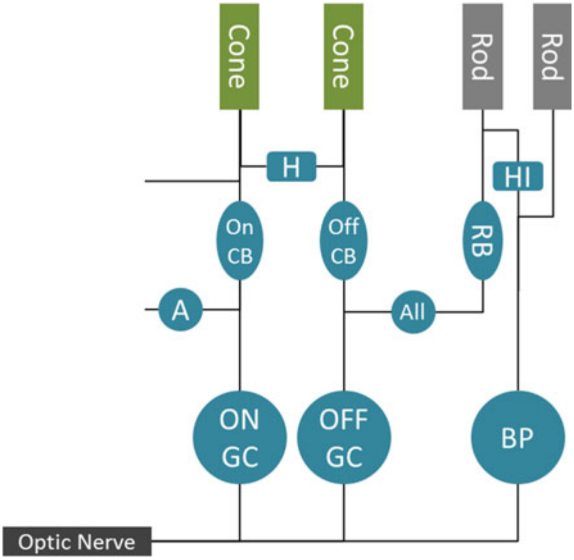


Fig. 5 Retinal interconnection of rods and cones via rod bipolars (RB), ON and OFF cone bipolars (CB), horizontal cells (H, HI), and amacrine cells (All, A) to ganglion cells (GC). Small dots indicate inhibitory, large dots excitatory connections. The horizontal interconnections as well as the connection of ON and OFF cells with cones show the formation of differential signals, whereas the rod signal is mainly superimposed onto the feed of the optical nerve

Ganglion Cells

Ganglion cells are classified by their anatomical properties: biplexiform, bistratified, midget, and parasol cells have been identified. Except the biplexiform cells, all types are crucial for higher-level processing in the visual system in respect of spatial, temporal, and color perception. Ganglion cells interface with multiple bipolar and amacrine cells in their surroundings forming a receptive field. This field defines a group of photoreceptors in the retina which influence the neuronal output of the connected ganglion cell.

Midget and parasol cells have a property known as spatial and spectral opponency which describes that ganglion cells distinguish between center and surrounding stimuli in their receptive field in respect of location and color. ON ganglion cells have an excitatory center and an inhibitory surround and respond stronger to bright spots (see Fig. 6). OFF ganglion cells respond best to dark spots. This property of ganglion cells can be seen as the second stage of a high-pass filtering (first stage are the horizontal cells) causing a higher sensitivity of the visual system for details (edges, spots, highlights, etc.) or high frequencies.

Midget cells (70–80 %) enable the visual system to perceive red-green color as well as dark-light luminance opponency. Blue-yellow opponency is sensed by bistratified ganglions which have a much lower incidence (<10 %). The ratio of red-green versus blue-yellow opponent ganglion cells reflects the high neuronal

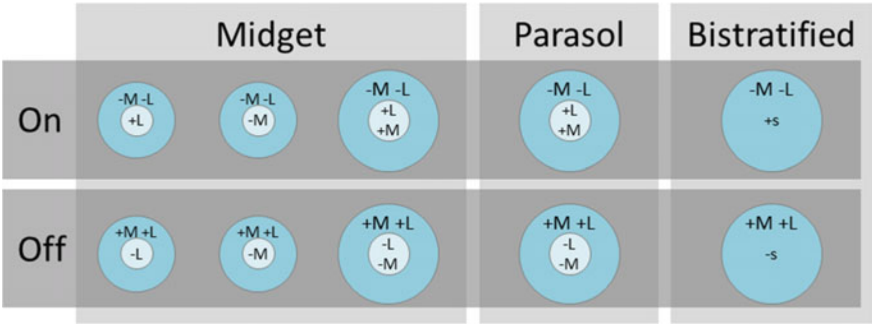


Fig. 6 Spatial and spectral receptive fields. Midget and parasol cells generate spatial and spectral differential signals between red (*L*) and green (*M*) cones, whereas bistratified cells interface blue (*S*) and red/green cones and show spectral but no spatial opponency

response and increased sensitivity for red-green color stimuli. Midget cells have a much smaller receptive field than parasol cells. Actually, midget cells in the fovea account only for a single cone. Therefore, the spatial resolution in the fovea is increased, which elevates the sensitivity for small scale details and color changes. However, parasol cells have larger receptive fields and cannot distinguish between colors but are sensitive to contrasts.

Thus, based on the size of their individual receptive fields, midget and parasol cells can be seen as high-pass filters sensitive to different frequencies or contrast gradients of different sizes in the visual field.

The processing in the inner nuclear layer of the retina described above are mainly designed to separate neuronal activity contrasts originating from the different cone types into luminance and color (red versus green and blue versus yellow) channels. This physiological finding forms the basis for perceptual color spaces which are able to specifically address the perceived lumiance contrast or color appearance.

To increase the spectral and spatial sensitivity, the HVS generates and amplifies differential signals in all processing stages. Additionally, the increase in size of the receptive fields and the overlay of low light sensitive monochromatic rod signals form a low-level compression stage which allows the HVS to transmit the information gathered by ≈120 mio photoreceptors via only ≈1 mio fibers in the optical nerve.

The fibers of the ganglion cells form the optical nerve conveying the compressed neuronal signals to the visual cortex via the visual pathway. However, before the signal reaches the cortex, a second preprocessing takes place in the lateral geniculate nuclei, which plays a large role in subconscious image processing.

3.2 The Lateral Geniculate Nuclei

The fibers of the optic nerve route into six different regions in the brain. The majority of fibers projects onto the LGN whereas superior colliculi, suprachiasmatic nuclei, pretectum, pregniculate, and the accessory optic system are secondary target areas. The LGN is connected to the striate cortex, the first stage of the visual cortex, and can be seen as the pre-processing stage for high-level image processing.

The secondary target regions of the optical nerve are thought to be responsible for unconscious visual processing such as synchronization of the biological clock, regulation of pupil diameter, (unconscious) direction of visual attention, and stabilization of the retinal image during head movement [11]. The most remarkable property of the LGN is its layered structure: Midget, parasol and bisratified ganglion cells project to different layers recreating a map of the entire visual field. Although the terminals are in different layers, the spatial locations of the originating photoreceptors are in register. Thus, if a spatial confined stimulus triggers the photoreceptors in the retina, neuronal activity will be triggered in all layers at nearly the same relative position. This organization reflects the separation of the perceived image into multiple representations of spatial resolution and color opponency at multiple scales. This separation allows to address single perceptual properties such as luminosity, color, or a specific spatial contrast frequency without interfering with other representations. This is utilized by the introduced rendering techniques to manipulate the human depth assessment in 2D images (see Chap. “Intuitive Motion and Depth Visualization for Rear-View Camera Applications”, Sects. 2 and 3).

Not all functions of the LGN are entirely understood. However, it is thought to mainly store the image representations and to sharpen spatial responses by enabling cross comparisons between different image representations. This enhances the salience of differential signals in the neuronal output emerged from the retina. Therefore, it can be seen as the third stage of image filtering which is comparable to a physiological implementation of multilateral image filtering.

3.3 The Visual Cortex

The visual cortex receives the low-level pre-processed differential signals from the LGN and generates a mental representation of the currently viewed scene. The receptive fields of the visual cortex are much more complex and show neuronal sensitivities for depth and motion cues as well as for temporal progression.

The first layer of the visual cortex is the striate cortex (V1). About half of its cells are devoted to the central 600 arcmin of the visual field (equals 1 % of the entire visual field [2, p. 201]). This is consistent with the higher amount of ganglion cells devoted to the central receptive field, interfacing the photoreceptors in the fovea centralis. Thus, spatial and color resolution for this part of the visual field are at its highest. The striate cortex has a high density of interconnections to the cortex and the extrastriate cortex. This enables the striate cortex to show new forms of selective stimulus sensitivity, particularly regarding spatial orientation, motion direction, binocularity, and wavelength. Hence, the striate cortex combines the neuronal information from different LGN-representations into larger and more complex spatial and temporal sensitive receptive fields.

The remaining visual areas are summarized under the term existriate cortex which is largely dependent on the processed neuronal output of V1. Its detailed functional significance for conscious visual processing is complex and still unclear. However, it is shown that the concept of expanding size and complexity of receptive fields continues in the extrastriate cortex [2, p. 202].

Orientation Tuning

Like all other optical systems, the HVS is limited in resolving high frequencies or details due to diffraction and aberrations. This limitation confines the human eye to a theoretical maximum resolution of about 0.6 arcmin of the visual field. This corresponds to the minimum cone spacing in the fovea (see Sect. 3.1). Although studies have shown that the visual system is able to resolve spatial details with this spacing, the decline in sensitivity at high and low frequencies is much steeper to be accountable to the optic degradation (see Fig. 7).

This degradation points to aliasing during the visual processing: frequency selective excitatory and inhibitory neurons are stimulated by specific detail gratings in their receptive fields. If there are less neurons devoted to a specific frequency, the HVS is less sensitive for this frequency. Thus, one can conclude that most neurons are sensitive to gratings at 10 arcmin (see Fig. 7) whereas the number of neurons sensitive to higher or lower frequencies is lower, resulting in the reversed u-shape of the human contrast sensitivity.

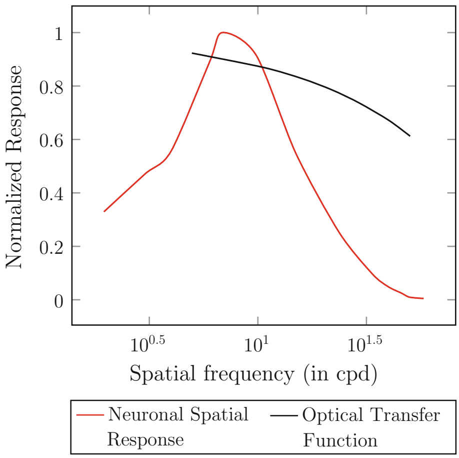


Fig. 7 Normalized human contrast sensitivity (Data from Campbell and Robson [12]) and normalized optical transfer function with a pupil diameter of 2 mm (Data from Ijspeert et al. [13]). The decline in human spatial sensitivity is much steeper than the optical decline. Thus, spatial sensitivity is driven by neuronal processing capacities

However, the neurons of the LGN do not show such a fine grating of frequency sensitivity [14]. Hence, decomposition of fine gratings is performed in the striate cortex: Orientation tuned neurons are sensitive to contrast edges of a very specific orientation and size. Their cellular structure is similar to ganglion and LGN cells: they emit neuronal signals if an edge with a specific orientation is sensed in their receptive field [15]. However, their receptive fields are even more complex and show narrow frequency band pass characteristics for various frequencies [16].

In conclusion, to specifically address the neural output induced by small and large scale gratings, the sensed input stimuli (e.g. an image) has to be decomposed into a similar representation of different spatial frequency scales first. Based on this decomposition, modulations of the stimuli frequency components can be utilized to affect the neural output for the neurons sensitive to specific spatial frequencies.

Motion Direction Selectivity

Besides spatial selectivity tuned neurons, there are also temporal selective cells which are sensitive to movement of edges in specific directions. Their neuronal activity is highest if an object is moved in the preferred direction. They do not respond if the edge is moved in the opposite or other directions [2, p. 199, 17]. A more detailed review of motion perception is given in Sect. 4.

Binocularity

Due to partial decussation of the optic nerve in the optical chiasm, fibers from the left and right visual fields of both eyes terminate in different hemispheres. Specialized binocular cortical cells compare the relative feature locations in the visual fields from both eyes. Their neuronal activity is highest if the retinal images of an object of left and right eye do not lie on corresponding retinal points but diverge in a specific angle. Thus the emitted neuronal signal represents the binocular disparity which encodes the distance to the considered object (see Sect. 7). This property of lateral shifted image capturing systems is utilized for stereo reconstruction in machine vision in a similar fashion [18].

Wavelength

The retinal ganglion cells already encode chromatic opponency in their neuronal signals. These signals are decoded and processed by the striate cortex. The responses of 41 % of the striate cells [19] show a chromatic opponency as well. This indicates sensitivity for wavelength variations and therefore color processing in the visual cortex.

The specific color and orientation tuning of neurons underpins the separation of color and edge information paths in the HVS, which motivates the separation for color and luminance in image manipulation in luminance contrast and hue.

3.4 Summary

The last section evaluates the human eye as the sensory organ for sight and the low-level processing of visual stimuli induced by the light inciding onto the retina. Human visual perception operates via a spatial and property based *divide and conquer* principle to assess the received optical stimulus. Already in the retina the signal is divided into color and luminance. Furthermore, the separated representations are transmitted as differential signals generated by neurons with a spatial confined receptive field. While ascending the optical pathway, these receptive fields grow in size and complexity, which allows to asses binocular disparity, motion, and shading on the highest stage (visual cortex). However, throughout the entire chain the specific properties are processed in separated pathways and are not merged until the final composition of the mental scene representation. This causes optical illusions since the stimulus is not assessed as a whole but as local contrasts (see Fig. 8). These processing paths can be specifically addressed and manipulated to influence the mental scene representation with respect to *color*, *disparity*, *motion*, *sharpness*, or *shading* perception.

Color is perceived by interpreting differential signals generated by wavelength sensitive cones. Due to the high amount of L-cones, the highest neuronal activation is induced by red colors.

Sharpness and *shading* are assessed by the secondary luminance contrast pathway. Besides the cones' spacing, this pathway defines the spatial resolution and

Fig. 8 *Simultaneous Contrast Illusion* shows the local spatial delimitation of perceptive fields. Although the central bar has a homogeneous color, it appears to fade from *dark* to *light grey*. This is induced by the local receptive field which only evaluates the local contrast between the bar and the background, the global contrast cannot be perceived

contrast sensitivity of the HVS. It decomposes the input image into several spatial frequency representations which encode high frequencies in the range of the cone spacing (sharpness) down to lower frequencies (e.g. shading cues).

Binocularity is assessed by neurons in the visual cortex which are sensitive to a lateral shift in the visual field. Thus, to address this processing path, different images have to be presented to the left and right eye (see Sect. 7.1).

Motion perception relies on luminance contrast decomposition of the visual pathway as well. Specifically tuned neurons in the visual cortex are sensitive to edges moving in a defined direction. Experimental findings indicate that this process does not necessarily require apparent motion [17]. This emerges the question if these neuronal structures can be stimulated by still images.

4 Motion Blur in Human Vision

The following section evaluates the impact of motion blur on human motion perception. The assessment of movement in the HVS can be separated into the physiological limitations resulting in motion blur and the neuronal processing of this artifact.

4.1 Inertia of Photoisomerization

As stated in Sect. 3.1, the process of photoisomerisation is not instant and not temporal confined. A stimulus with a brief flash of light takes up to 10 ms to generate a neuronal response and the neuronal activity lasts at least 100 ms for cones and up to 500 ms for rods (see Fig. 9).

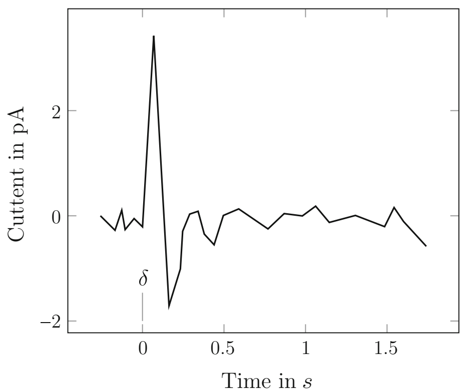


Fig. 9 Time course of macaque monkey photoreceptor current after a light impuls δ at $t = 0$ (Data from Schnapf and Baylor [20])

(a) (b)

London Bus 200 ms Waterfall 600 ms

Fig. 10 **a** Example of motion blur with a exposure time of 200 ms (from [21]). **b** Example of motion blur with a exposure time of 600 ms

The inertia of the chemical process of photoisomerization generates similar effects as long exposure times in cameras: If a bright light source passes along the static visual field, photoisomerization in multiple photoreceptors is induced. Since this process is not instant, neuronal activity is triggered along the trajectory of the bright object (e.g. A fast bright moving spot in a dark environment).

Therefore, the visual system has developed the mechanism of optokinetic nystagmus which is responsible for the tracking of focused moving objects along its trajectory. This makes motion blur for objects inside the focus of attention unperceivable.

However, this mechanism fails for high velocities (e.g. a spinning wheel or vent, fast vehicles) (see Fig. 10) and gets more apparent for moving bright light sources under low-light conditions. The latter is due to the increased receptor sensitivity to compensate for the low scene illumination.

The expansion of neuronal activity across several photoreceptors can be seen as a biological motion blur. Burr et al. [22] have determined exposure times of approximately 120 ms for cameras to achieve a blur similar to its biological equivalent.

While the physical impact of underlying chemical processes is quite clear, the way the visual system copes with and analyzes this information is still challenged in research. The main question is if motion cues in still images are able to activate motion-sensitive neurons and hence induce the subconscious perception of movement.

4.2 Neuronal Principles of Motion Sensation

On a lower level, motion is sensed by direction-selective neurons (see Sect. 3.3). To stimulate these neurons, an actual motion is required which is obviously not present in still images. However, research results indicate a fusion of temporal and spatial processing in the higher stages of the HVS: Burr and Ross [23] as well as Geisler [24] have found motion streaks arising from motion blur are subconsciously processed as time-independent shapes (i.e., horizontal lines, or “speedlines”) and then fused with motion perception to derive the apparent motion direction. Thus, motion blur depicted in a photographic still image triggers the same neuronal activity patterns as actual motion. Neuroimaging studies confirmed these findings [25–28]. The presentation of pictures with and without implied or depicted motion as well as physically motion led to activation of the same visual cortex regions as real motion.

These findings are underpinned by Winawer et al. [29] who analyzed the perceptual impact of motion depiction in still images on low-level neuronal processing. They investigated the MAE (Motion After Effect) which describes the effect that static displays appear to move in one direction if the viewer was previously preconditioned with a prolonged stimuli of a movement in the opposite direction [30]. This phenomenon results from differential signaling and inertia of neuronal processes described in Sect. 3. Winawer et al. showed that the MAE also occurs for static depiction of motion in still images. Therefore, the same motion-sensitive neurons are activated by still images showing motion cues as well as real motion. Burr et al. [23] support these findings by showing that even unnatural or exaggerated motion depiction expressed by speed lines has an impact on human motion perception.

4.3 Sensation of Movement in Still Images

In summary, neurological as well as perceptual studies have found evidence for natural and exaggerated motion blur to induce the perception of motion—independent form the apparent object motion. Thus, motion depictions in still images as well as in video sequences are a powerful, yet natural tool to convey additional motion information to the viewer.

5 Feature and Object Perception

The previous sections discussed the low-level decomposition of the visual stimulus in the retina and how it is compressed and pre-processed on its way to the striate cortex. In the following, models for the re-composition of these neuronal signals into combined scene representations are presented. Supportive renderings based on these models of human object perception, which are able to improve scene analysis which is shown in [1].

5.1 Feature Extraction

The visual system applies neuronal filters on the sensed image to decompose the stimulus into features and regions of different spatial frequencies, colors, and orientations. Multiple theories exist to describe the process of extracting image features based on the spatial filter output. They differ regarding the assumed stage of the visual pipeline in which the neuronal filter output is re-combined. However, they all agree that the visual system relies on edge information (orientation, spatial frequency, etc.) to generate a feature map of the image [31]. The combination of spatial frequency and orientation of edges is as well-known as texture. Landy and Graham [32] have shown that the visual system decomposes the image based on textures on a higher level of visual processing. Thus, the resulting feature map is built up of texture and color differences as well as on positions of edges including their orientation.

These findings describe the aggregation of neuronal outputs to features. The saliency of distinctive features creates a local elevation of excitatory levels inside the feature map, which leads visual attention to be drawn to these specific regions. The influence on visual is described in Sect. 6. The last transformation merges the features to particular texture-based object representations which form the building blocks for the final mental scene representation.

5.2 Image Segmentation

Texture separation based on feature extraction associates the physiological findings with the perceptual top-down theories of object depiction in the mental representation of the perceived scenery. To build up an object representation, its underlying shapes have to be recognized first. In the early 20th century the *Gestalt Laws* of Wertheimer [33] initially described the integration of features into complex shapes. As properties for the grouping of features into shapes proximity, similarity of color, similarity of size, common fate, and good continuation were stated (see Fig. 11).

Physiological findings support these assumptions. Features close to each other activate different groups of spatial filters than features which are more apart. These differences in filter output lead to a segmentation of the mental representation on a neuronal level. Similarity in size, orientation, and color of features can be explained in a similar fashion: Neuronal filter responses vary between the regions, forming a shape based segmentation for the next processing layer. Common fate describes the segmentation based on similar motion vectors or distance, which address groups of motion directed or spatial neuronal filters (see Sect. 3.3). Subsequently, these neuronal filter outputs are aggregated on a higher level to separate image regions.

Hence, the output of color, texture, depth, and motion selective filters form a neuronal image of the perceived scene. Higher level processing uses these attributes to segment the neuronal image representation [2, p. 251]. This supports the Gestalt

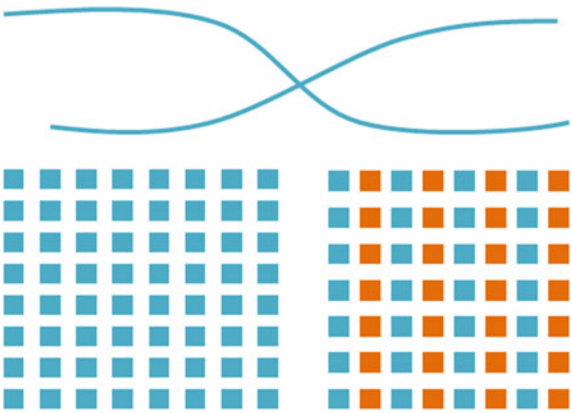


Fig. 11 Examples of the *Gestalt laws* (Proximity/similarity of color, Similarity of size and good continuation

laws of proximity, similarity, and common fate. Good continuation is supported by the model of collector units proposed by Morgan and Hotopf [34]. These units subsume responses of local orientation-selective neurons, whose preferred orientations slightly deviate, to form a smooth contour.

Several other theories for the formation of object representation exist [35–38], however, they all share the assumption that lower order features like depth, texture, or motion are integrated on a higher neuronal level to form a smooth contour, and that missing links are interpolated to form a closed silhouette.

5.3 Object Representation

The higher the ladder of visual processing is climbed up, the more complex the computations become. Somewhere between receptive field-based feature extraction and mental object representation the visual processing has to switch from image-based to a symbolic representation. There are several empirical theories about where and when this shift occurs [39–41]. Since the manipulation techniques described in the following are founded on lower feature map levels, higher human visual processing will not be taken further into account.

5.4 From Object Perception to Mental Scene Representation

The concept of neuronal features as the building blocks of the mental scene representation describes an accumulation of groups of tuned neurons (frequency, orientation or movement) responding to the stimuli inside the visual field. Based on these features the sensed image is segmented to form higher level object representations. Those are finally aggregated to the final mental scene representation.

The findings generated with the bottom-up approach of modern neurological science connect the early top-down theories of perception from the beginning of the 20th century. Both scientific approaches provide meaningful explanations for the visual system analysis of the viewed scene. Therefore, techniques can be derived to support the human scene understanding by simplifying the task of segmentation of different layers and objects. Additionally, this gives hints about how human visual attention is drawn to specific image regions and how the neuronal image representation is involved in the process of visual search-and-find.

6 Visual Search-and-Find

Besides supporting the HVS in accessing a given scene, guiding visual attention towards specific regions may support the identification of critical items or situations. Thus, speeding up the detection of a critical object may be utilized to support the viewer in search-and-find tasks.

Visual search-and-find tasks strongly rely on the mechanisms for filtering visual stimuli and the generation of mental scene representation. Accordingly, the findings in physiological science underpins the physiological model of the *Feature Interation Theory* of visual search by Treisman and Gelade [42] and its extension to *Guided Search* by Wolfe et al. [43].

6.1 Feature Integration Theory

As pointed out in the last section, the physiological characterization of visual processing is a bottom-up analysis of the human perceptual path of vision. Similar to the early theories of the *Gestalt Laws*, the perceptual research has approached the human visual processing in a top-down analysis. This analysis is performed by evaluating response times in search-and-find tasks. Treisman and Gelade [42] developed the *Feature Integration Theory* which separates visual search in a pre-attentive *Feature Search* and a attention-demanding processing stage *Conjunction Search*.

Feature Search is performed parallel and enables to immediately identify an object that distinguishes itself from the distractor objects in one distinct feature (e.g. a red circle in a field of blue circles). Treisman and Gelade's studies identified several kinds of such distinctive features: color, intensity, direction of lighting, orientation, size, motion direction, and disparity. The response times for feature search is below 200 ms and independent of the number of distractors. The latter indicates a parallel search, whereas the low response times support the assumption of unconscious processing, which directly highlights the searched object (see Fig. 12).

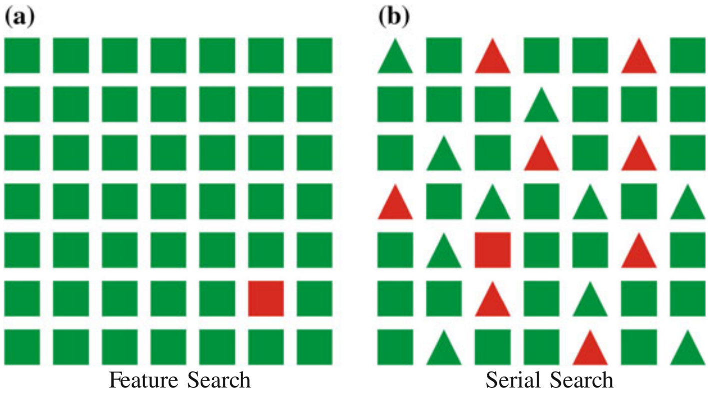


Fig. 12 **a** Examples for color popout-effect. **b** Example for conjunction-search where shape and color have to be combined. Following the *Guided Search Model*, *red* objects are searched first

This assumption is underpinned by the physiological findings (see Sect. 3.3) of tuned neuronal filters: Objects of similar depth, orientation, and spatial frequency induce the same neuronal output, generating a homogenous neuronal texture in the mental representation. An object located among these similar (distracting) objects, inducing a different neuronal output, causes a perturbance in the homogenous neuronal texture. This perturbance induces a singular neuronal activation of a distinct group of neurons. Thus, an object with different orientation, depth, color, or motion among a number of similar objects disrupts the homogenous neuronal representation of the perceived scene and causes higher level processing filters to automatically highlight the location of the perturbance in the visual field. Therefore, the position of objects with distinctive features is subconsciously and automatically determined by the visual system, which explains the low response times for this kind of search.

Additionally, Treisman and Gelade state that the serial *Conjunction Search* is performed if two or more features have to be combined for target identification (e.g. a red square in a field of red and green triangles and squares). The search requires attention and effort. Due to the serial processing of each displayed item, this search is much slower and response times increases with the number of distractors. This indicates that the features have to be combined on a higher neurological level, which requires conscious processing and therefore a longer processing time.

In terms of physiological research, the neuronal texture of the perceived scene is no longer homogeneous and the target is not highlighted due to the neuronal filter output. This explains the requirement for an item per item analysis by the visual system, leading to an exhaustive search.

6.2 Guided Search

Wolfe et al. [43] extended and modified the *Feature Integration Theory* proposing that visual search is executed in a subconscious bottom-up and a conscious top-down process at the same time. During the bottom-up analysis, all salient items are subconsciously pushed into the focus of attention (similar to the conjunction search for a single distinctive feature). A second conscious top-down analysis steers the viewer's attention towards the most promising features by applying a prior to the search.

For example, to find a blue pant inside a white closet, the top-down prior for steering attention generates higher activation levels for pants and blue items. The bottom-up analysis shows high activation for all colored and irregular shaped (clothes) items sorted in a descending order of their incidence. The combination of both analyses results in a ranking of items which define the order of attention focusing during search.

In terms of the example, an orange pullover might lead to a high activation in the bottom-up analysis but the prior from the conscious top-down process dampens this activation. Thus, in contrast to the *Feature Integration Theory* this search is not exhaustive since the objects are examined in a descending order.

This theory is underpinned by the physiological findings introduced in Sect. 5.3: neuronal activation to orientation, spatial frequency and color is filtered and sorted in a superordinate neuronal process to collect the most promising items in several neuronal image representations or *saliency maps*. Each representation encodes e.g. orientation, color or intensity. A higher-level perceptual process applies a prior to the neuronal representations and generates the final activation map [44, 45]. Subsequently, regions of highest activation are scanned first if they match the object to be found.

For pre-attentive search, the theories of *Guided Search* and *Feature Integration Theory* agree, since the highest activation is automatically generated by the most salient object. However, their assumptions are contradictory since the combination of bottom-up and top-down search allows an optimized search which does not require to be exhaustive. Physiological findings support the *Guided Search* model: The process of generating neuronal representations of the viewed scene on the basis of tuned neurons indicates a bottom-up analysis (see Sect. 5.1). On the other hand, there is no clear physiological evidence that higher processing stages are able to apply selective amplifications on specific features (e.g. a single color). However, perceptual studies [43–45] show a correlation between top-down search prior and search performance. Hence, an increased saliency can be used to push target objects into focus of attention to improve human search-and-find performance.

6.3 Speeding up Visual Search-and-Find

The unconscious bottom-up process is involuntarily driven by feature salience. Thus, manipulating salience to steer visual attention towards specific image regions is a well-known and commonly used technique in arts and rendering: Coarse brushwork leads to low spatial frequencies which are ranked low in the activation maps, whereas high frequencies induced by fine brushwork elevates the corresponding areas in the activation map [46]. A modern application of steering visual attention was presented by Cole et al. [47] who introduced a new rendering technique to manipulate contrast amplitudes (texture) and color in non-photorealistic renderings of architectural 3D-models. Dependent on the user-selected focus, the saliency (spatial frequency and color) is emphasized or dampened, causing the viewer's attention to be drawn into the specified image regions.

Kosara et al. [48] have shown that such a steering of visual attention is also possible by applying depth-of-field rendering to sharp input images. They unite the commonly used visualizations to steer visual attention known from arts and rendering and theories of visual search-and-find: Kosara et al. were able to show that renderings which draw the visual attention towards the target object shorten response time in search-and-find tasks [48].

Hence, by increasing the saliency of the target item, the viewer is able to find the target quicker. This could be achieved by adding a new distinctive feature (e.g. recoloring of the target or de-saturating the background) to the target item or by enhancing existing features (e.g. manipulating frequencies by enhancing or damping local contrast as well as introducing an artificial depth-of-field).

7 Human Depth Perception

For overall scene understanding, not only objects themselves but their spatial three-dimensional structure and arrangement is of high importance. Therefore, the HVS has developed multiple mechanisms to perceive depth and scene layering. By taking a close look on these mechanisms, it is very remarkable how much information is actually derived from two dimensional cues instead of binocularity. Hence, techniques for altering or inducing the perception of depth and distance by introducing or enhancing monocular depth cues are discussed.

7.1 Binocular Disparity

It is commonly assumed that human depth perception is mainly derived from the binocularity of the visual system. However, human binocular sight is strongly limited by its inter-ocular distance and resolution. In a closer analysis, the limits of depth resolution of this system are surprisingly low.

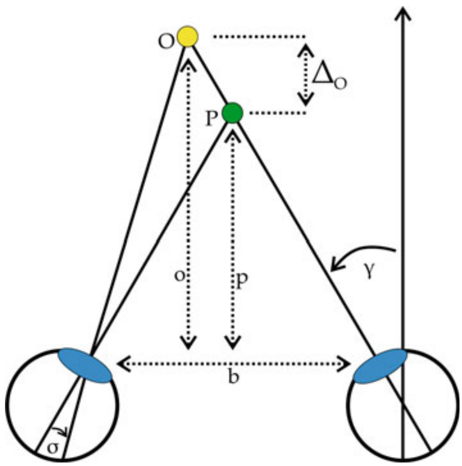


Fig. 13 Angular binocular disparity σ for the eyes baselength of b for objects in the distance o while fixating the point P, resulting in a shear angle of γ

Distance estimation from binocular disparity requires an offset in the relative retinal projection position of the same feature in both eyes. Figure 13 illustrates how objects at different depths manifest as retinal images with different disparities d on. The distance of fixation is p in which the relative disparity is zero ($\sigma = 0$). The disparity σ in angular units for the near and the far point depends on the base width of the stereo system $b \approx 65\,\mathrm{mm}$ and the distance to the point of fixation P as well as on the object distance o. To calculate the disparity σ, first the shear angle γ has to be calculated.

$$\gamma = \arctan\left(\frac{b}{2 \cdot p}\right) \tag{1}$$

$$\sigma = \gamma - \arctan\left(\frac{b}{d_O} - \frac{b}{2 \cdot p}\right) \tag{2}$$

For large distances, the disparity gets smaller until it is too small to fall on different relative photoreceptor locations in both eyes. Thus, the theoretical limit of resolving disparity is defined by the minimum photoreceptor spacing of 0.6 arcmin (see Sect. 3.1).

Accordingly, the visual system is able to detect depth variations Δ_O of at least 10 % of the fixation distance up to a distance of P = 33, 85 m (resulting in a resolution of 3.39 m per photoreceptor spacing). If the depth resolution is to be below 1 %, the distance of fixation p has to be smaller than 3.8 m to gain a minimum resolution of 38 cm in depth per photoreceptor spacing. As reference, it shall be noted that the maximum resolution at arm's length (p = 1 m) is about 2.7 mm per photoreceptor spacing (see Fig. 14). Therefore, binocular spatial resolution in depth decreases rapidly with growing distance: from millimeters at arm's length to meters in traffic situations.

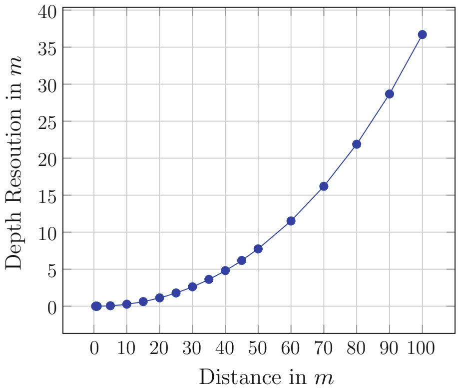


Fig. 14 Maximum achievable depth resolution (with 1 photoreceptor spacing) at various viewing distances in [m]. Smaller depth differences cannot be resolved. The resolution decreases rapidly with viewing distance

These limits are derived from the assumption of a perfect ability of the neurons in the visual cortex to process and correlate the features in each hemisphere to compute disparity. Howard and Rogers [49] have confirmed the minimum neuronal stereo processing capability of a resolution down to 0.6 arcmin. However, studies by Glennester [50] have shown that the maximum detectable disparity is 20 arcmin which equals about 33 photoreceptor spacings. Hence, the depth cue of binocularity is specialized for short distances <5–10 m. Similar results to these theoretical considerations have been verified by experimental approaches from Blakemore [51].

Furthermore, studies revealed that 2.7 % [52] of the general population is stereo-blind. Surprisingly, most of these people do not realize their limitation until they are confronted with artificial disparity stimuli (e.g. stereograms or 3D cinema). These findings arise the question why humans still have a strong sensation of depth and distance even if the HVS has reached the physiological limits of its stereo system or is not able to process stereo cues.

7.2 Monocular Depth Perception

The reason for strong depth and distance perception despite missing or limited stereo cues is the capability of the HVS to process additional pictorial as well as dynamic monocular depth cues, which are embedded in the 2D projection of the sensed scene. *Dynamic cues* arise from apparent motion of the visible objects or ego-motion (*motion parallax*) as well as the orientation of the eyes (*vergence*) or focused distance (*accommodation*). Dynamic cues will not be considered in detail since they do not suit to be manipulated within still images or video sequences to

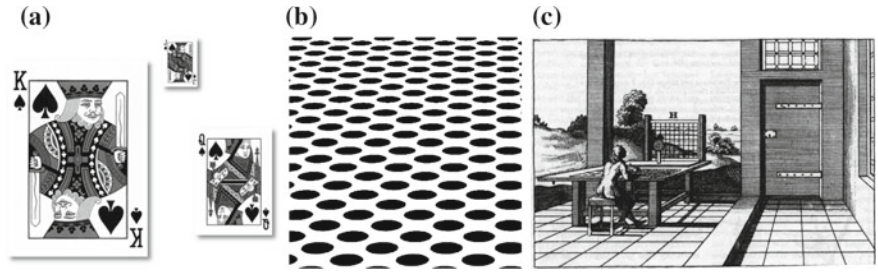


Fig. 15 **a** Relative size: the three playing cards appear to be located at different distances although they only differ in size. **b** Texture gradient shows the effect of foreshortening in a homogeneous texture. **c** Parallel perspective shows an apparatus for creating perspective drawings (Copperplate engraving from 1710)

enhance the perception of depth. However, *pictorial cues* encode, dependent on their origin, metric and ordinal depth information as well. They can be divided into two categories of predefined and modifiable cues.

Predefined cues are specified by the scenery and the object arrangement itself and cannot be altered, removed, or added in a trivial way to captured images of real scenes. Examples for these cues are relative *Retinal Image Size*, *Height in Visual Field*, *Parallel Perspective*, *Foreshortening* and *Interposition*.

Retinal Image Size arises from the distance-dependent scaling of object size on the retina. Naturally, for absolute distance estimation, the HVS requires either pre-knowledge of the size of the perceived object or a secondary reference object in the same field of view [53] (see Fig. 15a).

Height in the Visual Field is utilized by the HVS as a depth cue by evaluating the angle between a horizontal viewing direction and a point on the ground plane [54].

Parallel Perspective and *Foreshortening* are related to the cue of retinal image size and describe the scaling in width and the foreshortening of objects dependent on their distance (see Fig. 15c).

Parallel perspective is such a strong monocular cue that it induces a disturbed perception of object size in 2D drawings by giving the scenery an artificial impression of depth (see Fig. 16a). This is shown as well by the Ames room illusion, where the cue of *known size* and *relative size* are overruled.

The second category of pictorial cues is strongly coupled to the perceived spatial frequency (texture density gradient, blur from defocus) and color/luminance (atmospheric haze, shading). Therefore, these cues are directly processed on a low neuronal level in the HVS (see Sect. 3).

Since the perception of frequency and color/luminance can be manipulated easily by image processing algorithms, the cues presented in the following sections are suitable to manipulate the perceived depth in 2D images and videos.

Texture Density Gradient

The HVS analyzes distance-related foreshortening and scaling on the ground plane automatically resulting in a modifiable depth cue. If the ground plane has a homogenous texture, the density or spatial frequency of this texture increases with

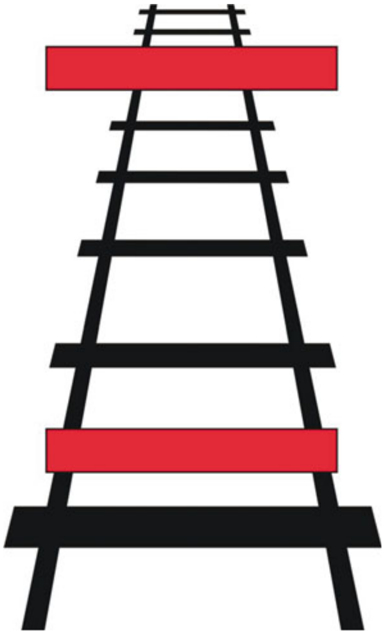


Fig. 16 The *Ponzo Illusion* induces the perception of distance due to parallel perspective; thus, the farther bar seems be larger although both bars have the same size. For both illusions, the cue of parallel perspective overrules the cue of relative or respectively known size

growing distance (see Fig. 15b). This cue is not only utilized by the HVS to assess distances but to estimate surface slant as well. Knill has studied this ability and was able to induce different perceptions of surface slant by manipulating the compression gradient [55].

This cue can be amplified or dampened by manipulating the texture's edge frequencies. To facilitate the perception of texture gradient, the contrast edges are enhanced to raise the output of frequency-tuned neurons. Thus, further processing stages for depth estimation from texture gradient receive stronger signals, which support their assessment. Accordingly, damping edge frequencies makes it harder for the visual system to evaluate the texture gradient. The impact on depth assessment gets apparent in Fig. 17.

Blur from Defocusing

The human eye only has a limited depth-of-field which is defined by aperture, focal distance, and resolution (see Sect. 2). Therefore, the eye is able to focus only at one distance at the same time whereas objects at farther or closer distances appear defocused or blurred. This retinal blur is evaluated subconsciously for the estimation of object distances in the visual field [56–58].

The importance and accuracy of blur from defocus as depth cue is still challenged by researchers till today. Some investigations have shown a contribution of blur to depth perception [56, 59] while others have found no [60] or only orderly effects on depth estimation [57, 61, 62]. Recent theories resolve these opponent findings by identifying a co-dependence of blur with other monocular cues (e.g. parallel perspective) to gather a correct depth estimation [5].

(a) (b) (c)

Dampened Texture Original Enhanced Texture

Fig. 17 The local contrast is increased from *left* to *right*. With growing neuronal stimulation at edges the cue of depth from texture gradient gets stronger. Therefore, the sensation of depth increases from *left* to *right* (**a** Bilateral filtered; **b** Original; **c** Unsharp masking)

Held et al. has confirmed this theory by developing "a probabilistic model of how viewers may use defocus blur in conjunction with other pictorial cues to estimate absolute distances to objects in a scene" [5], which has been validated with a perceptual study. Their findings indicate that blur is not only an ordinal but absolute depth cue if it is presented in conjunction with other monocular cues.

This conjuncture is utilized in photography and cinematography where cameras with a large aperture generate a confined depth-of-field to achieve two things: The conveyance of scene layering and inter-object distance information, and the separation of foreground and background objects. As depicted in Sect. 6, these pictorial cues are not only conveying depth information but are related to visual attention as well. An additional positive effect arises from the fact that images with proper defocus blur are generally perceived as more realistic and pleasant [63] (see Fig. 18).

(a) (b)

Circular Saw Husky

Fig. 18 Example of a photographs with a natural depth-of-field. **a** Shows a gradual blur gradient. **b** Shows an abrupt gradient. Shots taken by Anja Korte

Related work in artificial Blur from defocus

Therefore, the rendering of photorealistic blur from defocusing has been addressed early ($\sim$1981) in computer-generated imaging [64, 65]. Although processing power in computers has increased, depth-of-field rendering is still computational expensive. This limiting issue especially arises for real-time applications. To address this challenge, several faster approximation techniques have been proposed [66–69].

They tackle the problem by evaluating the importance of each pixel and apply high or low quality DOF (Depth Of Field) renderings. The adaption of these methods from computer imaging to photographs has been presented by Yu et al. [70]. Zhan et al. [71] introduced a real-time optimized implementation of the artificial depth-of-field synthesis proposed by Yu et al. to utilize this technique in video applications.

Shading as Depth Cue

Another strong monocular cue are shadows originating from object shape (*attached shadows*) and spatial arrangement (*cast shadows*). *Attached shadows* provide a powerful cue for perceiving 3D structures and shapes [72].

Even in the absence of texture the HVS is able to reconstruct the shape of an object solely from self-occlusion induced shading. When interpreting shading, the visual system tends to assume objects to be convex, not concave, and lighting from above. To support the perceptual analysis, the object has to have a uniform and diffuse surface illuminated by a diffuse and uniform light source. If these assumptions do not hold or the preconditions are not given, the interpretations based on shading alone are liable to be inaccurate.

However, by elevating the existing contrast between light and shadow, the perception of shading is facilitated. Thus local contrast enhancement on large and small scales is able to support human depth assessment (see Fig. 19).

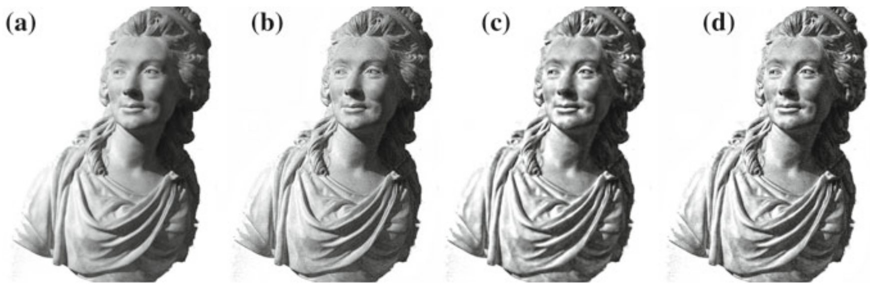


Fig. 19 Local contrast enhancement on small scales enhances the perception of texture of the material and fine structures (texture) whereas boosting on larger scales enhances the perception of shading. Combining both scales enhances the overall spatial sensation of the image. Image of *Comtesse de Sabran*; Photograph from Martin Dürrschnabel. **a** Original, **b** Boosting on small detail levels, **c** Boosting on large detail levels, **d** Boosting on all scales

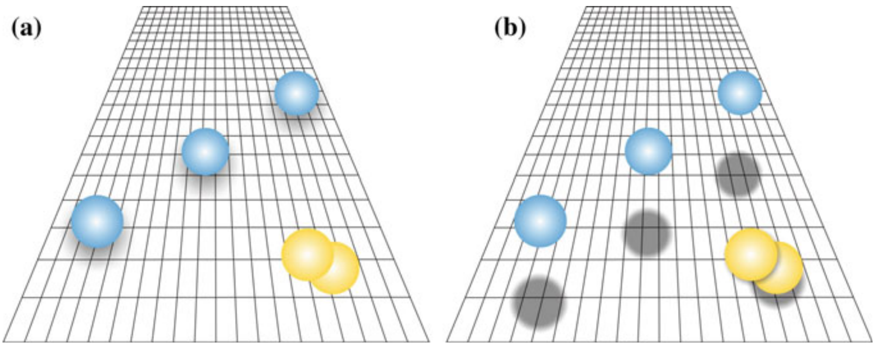


Fig. 20 The perceived 3D position of the *blue spheres* is affected by the position of the cast shadows. The cast shadows at the *yellow spheres* **a** original **b** show that shading supports the assessment of scene layering and foreground extraction

Kersten et al. [73] have shown that *cast shadows* originating from an object falling onto the surface are interpreted subconsciously to provide depth and layering information. This cue is as well suited to emphasize the perception of object interposition. By manipulating contrast, this ordinal depth cue can be weakened or enhanced, and by adding *cast shadows* it can even be artificially introduced (see Fig. 20).

Contrast manipulation is similar to the techniques for attached shadows whereas adding this cue artificially requires different approaches. Several algorithms have been introduced which add cast shadows to computer-generated renderings. *Ambient Occlusion* [74] is a well-known technique to provide a more realistic appearance for rendered images. The attenuation of light is calculated by casting rays in every direction at each position. If a ray intersects with the sky, the luminance at the origin is increased. The result is blended over the luminance channel of the output image. A faster screen space approximation technique has been presented by Luft et al. [75]. It is based on an unsharp mask of the depth buffer which is added to the luminance channel of the output image. Although the cast shading is physically not correct, the authors' findings indicate that users prefer renderings in comparison to the original images.

Saturation

A color-dependent cue arises from *atmospheric haze*. For large distances of several kilometers, the contrast of a viewed object decreases due to atmospheric scattering of light by floating particles (see Fig. 21). The human visual system takes this effect into account for an ordinal layering of the visible objects [76]. Additionally, Troscianko et al. have found evidence that saturation gradients are able to convey depth information in a similar fashion as texture gradients [77].

Thus, by depth controlled manipulation of color saturation in captured real images the sensation of depth gradient and scene layering is improved. Therefore,

Fig. 21 Example of atmospheric haze showing contrast attenuation due to scattering of light by atmospheric particles. Thus, more distant hills are rendered desaturated and with less contrast

such a rendering is proposed in Chap. 15, Sect. 3 to support the perception of scene layering and distance for an improved foreground and background separation.

7.3 Cue Combination

The previous sections show that all the described cues—binocular as well as monocular—contribute to human depth perception.

Based in these findings, it is indicted that all these cues are accumulated to one aggregated sensation of depth and distance. The accuracy of each cue differs depending on distance as well as on the quality and availability in the viewed scene. Therefore, it is reasonable to assume that the visual system encodes these dependencies in its neuronal structures. The cues are weighted according to their reliability before they are aggregated into a merged sensation of depth and distance. This assumption is supported by multiple empirical findings [78–80] which show a dynamic cue weighting of the HVS based on distance [81] and stimulus strength [79]. Due to the underlying empiric nature of the encoded weighting, cues which are usually counted to be reliable might overrule weaker cues even if they suggest an incorrect spatial structure—the result is commonly known as optical illusion (see Fig. 16).

On the other hand, if strong cues are not present (e.g. binocularity for 2D images), the visual system automatically evaluates the available (monocular) cues to gather the required distance information.

8 Conclusion

The previous analysis of the human visual system reveals the complex nature of the processes between the occurrence of a photon hit on the retina and an almost complete scene understanding in the cortical representation. Whereas the transformation of light signals to neuronal signals is very similar to the process taking place in a digital camera, already the first retinal processing stages are much more complex. The image is decomposed (e.g. into color and luminance) and compressed into various representations by generating differential signals of the sensed image. The resulting multi-scale representations gain complexity and abstraction with each processing stage. Hence, by modulating the sensory input (e.g. by manipulating the presented image material) for specific low-level neural structures the higher level mental scene representation can be altered. This yields an modified sensation of sharpness, color, luminance, depth, motion and saliency.

References

1. C. Rößing: Video and image manipulation for enhanced perception. Dissertation, University Ulm. http://vts.uni-ulm.de/doc.asp?id=9319 (2014)
2. Mather, G.: Foundations of Perception, 1st edn., p. 388. Psychology Press, Hove (2006). ISBN: 0863778348
3. Riordan-Eva, P., Cunningham, E.T.: Vaughan Asbury's General Ophthalmology, p. 492. Mcgraw-Hill Professional (2011). ISBN: 0071634207
4. Hess, E.H., Polt, J.M.: Pupil size as related to interest value of visual stimuli. Science **132**, 349–350 (1960). doi:10.1126/science.132.3423.349, ISSN: 0036-8075
5. Held, R.T., Cooper, E.A., O'Brien, J.F., Banks, M.S.: Using blur to affect perceived distance and size. ACM Trans. Graph. **29**(2), 1–16 (2010). doi:10.1145/1731047.1731057, http://portal.acm.org/citation.cfm?doid=1731047.1731057 ISSN: 07300301
6. Rushton, W.: Rhodopsin measurement and dark-adaptation in a subject deficient in cone vision. J. Physiol. **156,** 193–205 (1961)
7. Purves, D., Augustine, G.J., Fitzpatrick D.: Neuroscience, 3rd edn., pp. 1–773. Sinauer Associates, Sunderland (2004). doi: 978-0878937257, http://www.amazon.com/Neuroscience-CDROM-Dale-Purves/dp/0878937250, ISBN: 0878937250
8. Lennie, P.: Color vision: putting it together. Curr. Biol. **10**(16), R589–R591 (2000). doi:http://dx.doi.org/10.1016/S0960-9822(00)00632-1 . http://www.sciencedirect.com/science/article/pii/S0960982200006321, ISSN:0960-9822
9. Stockman, A., Sharpe, L.T.: The spectral sensitivities of the middle- and long-wavelength-sensitive cones derived from measurements in observers of known genotype. Vis. Res. **40**(13), 1711–1737 (2000). doi:10.1016/S0042-6989(00)00021-3, ISSN: 00426989
10. R. H. Masland, "The fundamental plan of the retina.", *Nature neuroscience*, vol. 4, pp. 877–886, 2001, issn: 1097-6256. doi: 10.1038/nn0901-877
11. R. W. Rodieck, *The First Steps in Seeing*. Sinauer, 1998, p. 367
12. Campbell, F.W., Robson J.G.: Application of fourier analysis to the visibility of gratings. J. Physiol. **197**, 551–566 (1968). ISSN: 0022-3751
13. IJspeert, J.K., van den Berg, T.J., Spekreijse, H.: An improved mathematical description of the foveal visual point spread function with parameters for age, pupil size and pigmentation. Vis. Res. **33**, 15–20 (1993). doi:10.1016/0042-6989(93)90053-Y, ISSN: 00426989

14. Derrington, A.M., Lennie, P.: Spatial and temporal contrast sensitivities of neurones in lateral geniculate nucleus of macaque. J. Physiol. **357**, 219–240 (1984)
15. Hubel, D.H., Wiesel, T.N.: Receptive fields and functional architecture of monkey striate cortex. J. Physiol. **195**(1), 215–243 (1968)
16. De Valois, R.L., Albrecht, D.G., Thorell L.G.: Spatial frequency selectivity of cells in macaque visual cortex. Vis. Res. **22**(5), 545–59 (1982). http://www.ncbi.nlm.nih.gov/pubmed/7112954, ISSN: 0042-6989
17. Mikami, A., Newsome, W.T., Wurtz, R.H.: Motion selectivity in macaque visual cortex. I. Mechanisms of direction and speed selectivity in extrastriate area MT. J. Neurophysio. **55** (6), 1308–1327 (1986). ISSN: 0022-3077
18. Hirschmüller, H., Innocent, P.R., Garibaldi, J.: Real-time correlation-based stereo vision with reduced border errors. Int. J. Comput. Vis. **47**(1), 229–246 (2002). doi:10.1023/A:1014554110407, http://www.springerlink.com/content/72d01vcdjd798bay/, ISSN: 09205691
19. Johnson, E.N., Hawken, M.J., Shapley, R.: The spatial transformation of color in the primary visual cortex of the macaque monkey. Nat. Neurosci. **4**(4), 409–16 (2001). doi: 10.1038/86061, http://www.ncbi.nlm.nih.gov/pubmed/11276232, ISSN: 1097-6256
20. Schnapf, J., Baylor, D.: How photoreceptor cells respond to light. Sci. Am. **254**(4), 32–39 (1987)
21. E01: London bus. http://flickr.com/photo/10158179@N06/2334039881. Accessed 03 Apr2014
22. Burr, D.C., Ross, J., Morrone, M.C.: Seeing objects in motion. In: Proceedings of the Royal Society of London. Series B, Containing papers of a Biological character. Royal Society (Great Britain), vol. 227(1247), pp. 249–265, 1986. doi:10.1098/rspb.1986.0022, ISSN: 0962-8452
23. Burr, D.C., Ross, J.: Direct evidence that “speedlines” influence motion mechanisms. J. Neurosci. (The official journal of the Society for Neuroscience) **22**(19), 8661–8664 (2002). doi:22/19/8661[pii], ISSN: 1529-2401
24. Geisler, W.S.: Motion streaks provide a spatial code for motion direction. Nature **400**(6739), 65–69 (1999). doi:10.1038/21886, ISSN: 0028-0836
25. Kourtzi, Z., Kanwisher, N.:Activation in human MT/MST by static images with implied motion. J. Cogn. Neurosci. **12**(1), 48–55 (2000). doi:10.1162/08989290051137594, http://dx.doi.org/10.1162/08989290051137594, ISSN: 0898-929X
26. Lorteije, J.A.M., Kenemans, J. L., Jellema, T., van der Lubbe, R.H.J., Lommers, M.W., van Wezel, R.J.A.: Adaptation to real motion reveals directionselective interactions between real and implied motion processing. J. Cogn. Neurosci. **19**(8), 1231–1240 (2007). doi:10.1162/jocn.2007.19.8.1231, ISSN: 0898-929X
27. Peuskens, H., Vanrie, J., Verfaillie, K., Orban, G.A.: Specificity of regions processing biological motion. Eur. J. Neurosci. **21**(10), 2864–2875 (2005). doi: 10.1111/j.1460-9568.2005.04106.x, ISSN: 0953-816X
28. Senior, C., Barnes, J., Giampietro, V., Simmons, A., Bullmore, E.T., Brammer, M., David, A.S.: The functional neuroanatomy of implicit-motion perception or representational momentum. Curr. Biol.: CB, **10**(1), 16–22 (2000) doi: 10.1016/S0960-9822(99)00259-6, ISSN: 09609822
29. Winawer, J., Huk, A.C., Boroditsky, L.: A motion aftereffect from still photographs depicting motion. Psychol. Sci.: A J. Am. Psychol. Soc./APS, **19**(3), 276–283 (2008). doi:10.1111/j.1467-9280.2008.02080.x, ISSN: 0956-7976
30. Wohlgemuth, A.: On the after-effect of seen movement. Ph.D. thesis, p. 148, Cambridge (1911)
31. Bruce, V., Green, P., Georgeson, M.: Visual Perception: Physiology, Psychology and Ecology, p. 496. Routledge (2003). ISBN: 1841692387
32. Landy, M.S., Graham, N.: Visual perception of texture. In: Chalupa, L.M., Werner, J.S. (eds.) The Visual Neurosciences, pp. 1106–1118. MIT Press, Cambridge (2004). ch. 9
33. Wertheimer, M.: Experimentelle studien über das sehen von bewegung. *Zeitschrift für Psychologie* (1912) http://www.getcited.org/cits/PP/1/PUB/103413947
34. Morgan, M.J., Hotopf, W.H.: Perceived diagonals in grids and lattices. Vis. Res. **29**(8), 1005–1015 (1989). doi:10.1016/0042-6989(89)90115-6, ISSN: 00426989
35. Marr, D.: Early processing of visual information. Phil. Trans. R. Soc. Lond. B Biol. Sci. **275** (942), 483–519 (1976). ISSN: 0962-8436

36. Nakayama, K., He, Z.J., Shimojo, S.: Visual surface representation: a critical link between lower-level and higher-level vision. In: Kosslyn, S., Osherson, D. (eds.) In Invitation to Cognitive Science, pp. 1–70. MIT Press, Cambridge (1995). http://books.google.com/books?hl=en, ISBN: 0-262-15042-5
37. Moulden, B.: Collator units: second-stage orientational filters. In: Ciba Foundation symposium, vol. 184, pp. 170–184. Discussion 184-192, 269-271 (1994)
38. Field, D.J., Hayes, A., Hess, R.F.: Contour integration by the human visual system: evidence for a local "association field". Vis. Res. **33**(2), 173–193 (1993). doi:10.1016/0042-6989(93)90156-Q, ISSN: 00426989
39. Biederman, I., Gerhardstein, P.C.: Recognizing depth-rotated objects: evidence and conditions for three-dimensional viewpoint invariance. J. Exp. Psychol. Hum. Percept. Perform. **19**(6), 1162–1182 (1993). doi:10.1037/h0090355, ISSN: 0096-1523
40. Tarr M.J., Bülthoff, H.H.: Is human object recognition better described by geon structural descriptions or by multiple views? comment on biederman and gerhardstein (1993). J. Exp. Psychol. Hum. Percept. Perform. **21**(6), 1494–1505 (1995). doi: 10.1037/0096-1523.21.6.1494, ISSN: 0096-1523
41. Marr, D., Nishihara, H.K.: Representation and recognition of the spatial organization of three-dimensional shapes. In: Proc. R. Soc. Lond.. Ser. B, Containing papers of a Biological character. Royal Society (Great Britain), 200(1140), 269–294 (1978). doi:10.1098/rspb.1978.0020, ISSN: 0962-8452
42. Treisman, A., Gelade, G.: A feature-integration theory of attention. Cogn. Psychol. **12**(1), 97–136 (1980). doi: 10.1016/0010-0285(80)90005-5, http://www.ncbi.nlm.nih.gov/pubmed/7351125, ISSN: 00100285
43. Wolfe, J.M.: Guided search 2.0: a revised model of visual search. Psychon. Bull. Rev. 1(2), 202–238 (1994). doi:10.1037/0096-1523.15.3.419. , http://www.springerlink.com/index/C0234T6313755617.pdf, ISSN: 10699384
44. Itti, L., Koch, C.: Computational modelling of visual attention. Nat. Rev. Neurosci. **2**(3), 194–203 (2001). doi:10.1038/35058500, ISSN: 1471-003X
45. Desimone, R., Duncan, J.: Neural mechanisms of selective visual attention. Ann. Rev. Neurosci. **18**, 193–222 (1995). doi:10.1146/annurev.ne.18.030195.001205, ISSN: 0147006X
46. Riebe, C., DiPaola, S., Enns, J.T.: Rembrandt's textural agency: a shared perspective in visual art and science. Leonardo (2010). doi:10.1162/leon.2010.43.2.145
47. Cole, F., Decarlo, D., Finkelstein A., Kin K., Morley K., Santella A.: Directing gaze in 3d models with stylized focus. Focus (2006)
48. Kosara, R., Tscheligi, M.: Useful properties of semantic depth of field for better f + c visualization, Image. Rochester, New York (2002)
49. Howard, I.P., Rogers, B.J.: Binocular vision and stereopsis, p. 736. Oxford University Press, Oxford (1995)
50. Glennerster, A.: Dmax for stereopsis and motion in random dot displays. Vis. Res. **38**(6), 925–935 (1998). doi:10.1016/S0042-6989(97)00213-7, ISSN: 00426989
51. Blakemore, C.: The range and scope of binocular depth discrimination in man. J. Physiol. **211**(3), 599–622 (1970)
52. Richards, W.: Stereopsis and stereoblindness. Exp.Brain Res. Experimentelle Hirnforschung. Experimentation cerebrale **10**(4), 380–388 (1970). doi:10.1007/BF02324765, ISSN: 0014-4819
53. Ittelson, W.H.: Size as a cue to distance; radial motion. Am. J. Psychol. **64**(2), 188–202 (1951)
54. Ooi, T.L., Wu, B., He, Z.J.: Distance determined by the angular declination below the horizon. Nature **414**(6860), 197–200 (2001). doi:10.1038/35102562, ISSN: 0028-0836
55. Knill, D.C.: Discrimination of planar surface slant from texture: human and ideal observers compared. Vis. Res. **38**(11), 1683–1711 (1998). doi:10.1016/S0042-6989(97)00325-8, ISSN: 00426989
56. Pentland, A.P.: A new sense for depth of field. IEEE Trans. Pattern Anal. Mach. Intell. **9**(4), 523–531 (1987). doi:10.1109/TPAMI.1987.4767940, ISSN: 0162-8828

57. Marshall, J.A., Burbeck, C.A, Ariely, D., Rolland, J.P., Martin K.E.: Occlusion edge blur: a cue to relative visual depth. J. Opt. Soc. Am. A, Opt., Image Sci. Vis. **13**(4), 681–688 (1996)
58. Mather, G.: Image blur as a pictorial depth cue. In: Proc. R. Soc. Lond. Biol. Sci. **263**, 169–172 (1996)
59. Watt, R.J., Morgan, M.J.: Spatial filters and the localization of luminance changes in human vision. Vis. Res. **24**(10), 1387–1397 (1984). doi:10.1016/0042-6989(84)90194-9, ISSN: 00426989
60. Mather, G., Smith, D.R.R.: Blur discrimination and its relation to blurmediated depth perception. Perception **31**(10), 1211–1219 (2002)
61. Mather, G.: The use of image blur as a depth cue. Perception **26**(9), 1147–1158 (1997)
62. Palmer, S.E., Brooks, J.L.: Edge-region grouping in figure-ground organization and depth perception. J. Exp. Psychol. Hum. Percept. Perform. **34**(6), 1353–1371 (2008)
63. Hillaire, S., Lecuyer, A., Cozot, R., Casiez, G.: Using an eye-tracking system to improve camera motions and depth-of-field blur effects in virtual environments. In: Proceedings of the Virtual Reality Conference, VR '08. IEEE, vol. 28(6), pp. 47–50 (2008). doi:10.1109/VR.2008.4480749, ISSN: 0272-1716
64. Potmesil, M., Chakravarty, I.: A lens and aperture camera model for synthetic image generation. ACM SIGGRAPH (1981). doi:10.1145/965161.806818
65. Porter, T., Carpenter, L., Cook, R.L.: Distributed ray tracing (1984). doi:10.1145/964965.808590
66. Fearing, P.: Importance ordering for real-time depth of field, in *Image Analysis Applications and Computer Graphics*, 1995, pp. 372–380, isbn: 978-3-540-60697-0
67. Rokita, P.: Generating depth-of-field effects in virtual reality applications. IEEE Comput. Graph. Appl. **16**(2), 18–21 (1996)
68. Barsky, B.A.: Vision-realistic rendering: simulation of the scanned foveal image from wavefront data of human subjects. In: Proceedings of the 1st Symposium on Applied perception in graphics and visualization - APGV '04, vol. 1, p. 73. ACM Press (2004). doi:10.1145/1012551.1012564, http://dl.acm.org/citation.cfm?id=1012564, ISBN: 1581139144
69. Mulder, J.D., van Liere R.: Fast perception-based depth of field rendering. In: Proceedings of the ACM Symposium on Virtual Reality Software and Technology, ser. VRST '00, pp. 129–133, New York (2000).doi:10.1145/502390.502414, http://doi.acm.org/10.1145/502390.502414, ISBN: 1-58113-316-2
70. Yu, X., Wang, R., Yu, J.: Real-time depth of field rendering via dynamic light field generation and filtering. Eecisudeledu **29(**7). Alliez, P., Bala, K., Zhou, K. (eds.) (2010). http://www.eecis.udel.edu/~jye/lab%5C_research/10/pgdof.pdf
71. Yu, Z., Thorpe, C., Yu, X., Grauer-Gray, S., Li, F., Yu J.: Dynamic depth of field on live video streams: a stereo solution. Cis.udel.edu. http://www.cis.udel.edu/~feli/%5C_papers/cgi11.pdf
72. Koenderink, J.J., van Doorn, A.J., Kappers, A.M.: Pictorial surface attitude and local depth comparisons. Percep. Psychophys. **58**(2), 163–173 (1996), doi:10.3758/BF03211873, ISSN: 0031-5117
73. Kersten, D., Knill, D.C., Mamassian, P., Bülthoff, I.: Illusory motion from shadows. (1996). doi:10.1038/379031a0
74. Langer, M.S., Bülthoff, H.H.: Depth discrimination from shading under diffuse lighting. Perception **29(**6), 649–660 (2000). http://www.ncbi.nlm.nih.gov/pubmed/11040949
75. Luft, T., Colditz, C., Deussen, O.: Image enhancement by unsharp masking the depth buffer (2005)
76. O'Shea, R.P., Blackburn, S.G., Ono, H.: Contrast as a depth cue. Vis. Res. **34(**12), 1595–1604 (1994). doi:10.1016/0042-6989(94)90116-3, ISSN: 00426989
77. Troscianko, T., Montagnon, R., Le Clerc, J., Malbert, E., Chanteau, P.L.: The role of colour as a monocular depth cue. Vis. Res. **31**(11), 1923–1929 (1991). http://www.ncbi.nlm.nih.gov/pubmed/1771776
78. Buelthoff, H., Mallot, H.: Integration of stereo, shading and texture. In: Blake and Troscianko (eds.) Al and the eye, pp. 119–146. Wiley, New York (1990)

79. Buckley, D., Frisby, J.P.: Interaction of stereo, texture and outline cues in the shape perception of three-dimensional ridges. Vis. Res. **33(**7), 919–933 (1993). . doi:10.1016/0042-6989(93)90075-8, ISSN: 00426989
80. Cumming, B.G., Johnston, E.B., Parker, A.J.: Vertical disparities and perception of three-dimensional shape. Nature **349**(6308), 411–413 (1991) doi:10.1038/349411a0, ISSN: 0028-0836
81. Johnston, E.B., Cumming, B.G., Landy, M.S.: Integration of stereopsis and motion shape cues. Vis. Res. **34**(17), 2259–2275 (1994) doi:10.1016/0042-6989(94)90106-6, ISSN: 00426989

Camera Monitor Systems Optimized on Human Cognition—Fundamentals of Optical Perception and Requirements for Mirror Replacements in Commercial Vehicles

Albert Zaindl

Abstract The way that fields of views are displayed is extremely important for finding a suitable replacement of the existing set of mirrors by Camera Monitor Systems (CMS). Especially in commercial vehicles, the great number of required mirrors (currently at least six pieces!) poses a great cognitive challenge to the driver. The smart method of image presentation in the CMS can provide marked relief to the driver and thus considerably improve safety. The key prerequisite for optimizing image presentation is to understand human perception and to derive therefrom the necessary requirements for the CMS. Before the driver turns their attention to an object, this object is normally detected in their peripheral field of view due to a change in optical flow. Current mirror systems for commercial vehicles violate this principle. The driver must specifically look at the mirror in order to be able to detect any objects inside. However, the possibility for peripheral discovery of objects at close ranges (wide-angle mirrors or close proximity mirrors) would be very beneficial and noticeably relieve the driver. It can be concluded that the single replacement of each individual mirror is not an expedient solution; instead, the focus should be placed on merging the different fields of view. The CMS of MAN Truck & Bus combines the fields of view from the main, wide-angle and close proximity mirrors and helps the driver keep track of their vehicle and the surrounding areas. Thus, the system improves safety by enabling drivers to see what is happening around them more quickly and accurately.

Keywords Foveal vision · Peripheral vison · Optical flow · Combined fields of view · Ergonomics · Safety · Commercial vehicles · Camera monitor system · Mirror replacement

A. Zaindl (✉)
Employee at MAN Truck & Bus AG & Doctoral Fellow at the Institute of Ergonomics, Technical University of Munich, Munich, Germany
e-mail: Zaindl@tum.de

A. Terzis (ed.), *Handbook of Camera Monitor Systems*,
Augmented Vision and Reality 5, DOI 10.1007/978-3-319-29611-1_10

Preamble *This paper represents a combination and extension of two articles that have already been published in the "ATZ—Automobiltechnische Zeitschrift"* [1] *and the "Ergonomie Aktuell"* [2]. *Parts of some sentences in this paper are copied verbatim from these two articles. For reasons of clarity, the respective segments are not quoted separately.*

1 Fields of View in Commercial Vehicles

Anyone who gets into a commercial vehicle for the first time will be amazed at the number of mirrors that are installed on the "King of the Road". All of these mirrors were introduced slowly over time in order to eliminate safety deficiencies. And while long-distance visibility in commercial vehicles can now be considered largely ideal due to the high seat position, the direct view of the surrounding and rear areas from inside the truck still leaves a lot to be desired. These otherwise non-visible areas around and behind the truck can only be viewed through mirrors. Current laws require at least six pieces. These comprise the two main side view mirrors (class II), the two wideangle side view mirrors (class IV), as well as a close-proximity exterior mirror on the passenger's side (class V) and a front view mirror (class VI). Figure 1 shows the fields of view for commercial vehicles as currently required by law [3]. Figure 2 shows the view of the driver via the passenger side mirror.

During the introduction of new field of view classes, care was taken to eliminate as many blind spots as possible. A further enlargement of the class V field of view for high vehicle cabs in 2015 enables the driver to maintain an overview of the entire side area of the truck [4]. This reduces the blind spot to a minimum.

In 2004, Spiegel Online ran the headline "The blind spot has been eliminated" (in German: "Der tote Winkel ist tot"). "From a visual perspective, the safety issue is now as good as resolved. However, just like before, truck drivers are still responsible for focusing their attention on all (…) mirrors before any change of direction" [6].

Fig. 1 Field of view classes for commercial vehicles (according to [1])

Fig. 2 View via the passenger side mirrors [5]

This paper addresses precisely the point in the article. The variety of mirrors and different fields of view imposes a task on drivers that is inherently difficult from a cognitive perspective: Drivers have to simultaneously monitor the road ahead and the 6 additional mirrors that are mandatory at present. Which mirror should the driver look at, and at which point, in order to correctly assess the road situation? With what frequency should they look at the mirrors and what is made visible by each mirror?

In order to better explain the human capacity of sight, this paper explores the fundamental characteristics of foveal and peripheral vision—whose combination constitutes human sight. The fundamental requirements for mirror replacement systems are stipulated on the basis of these characteristics and the related cognitive abilities.

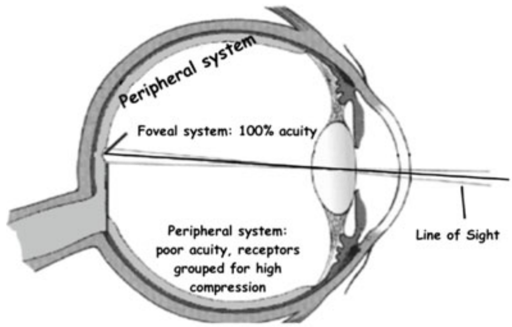
2 Foveal and Peripheral Vision: The Cornerstones of Human Sight

This section deals with human optical perception in general, with the focus being placed on the two different types of vision (foveal and peripheral vision). The relationship of these two types of vision with optical flow is used to explain the efficiency of human vision. This section represents the theoretical foundation of the chapter, and has been summarized from the relevant literature [7–10]. *The next sections expound on these fundamental insights but do not require them in order to be understood.*

Optical perception in humans can be divided into two different ways of seeing: foveal and peripheral vision. Both types combine to form a powerful overall construct that makes up human optical perception, i.e. human sight. Understanding both kinds of vision and their combined effectiveness is essential to comprehend the overall perception of the human eye and the associated cognitive processes.

The eye is an organ that converts optical signals into electrical ones. These electrical signals are then transmitted to the brain, which handles the image processing. In a sense, the human eye can be compared to a camera and the brain to a graphic processing unit. While a camera receives the light signals by means of so-called pixels, in human sight this process takes place by way of cones (photopic vision) and rods (scotopic vision), which are located on the retina. In contrast to cameras, these receptors are not evenly distributed on the retina. While there is a very high concentration of cones in the main direction of sight (foveal vision), rods are predominant in the periphery and are not found in the fovea. The general structure of the eye and the density of the photoreceptor cells on the retina over the viewing angle are shown schematically in Fig. 3.

The light from the central line of sight falls onto the fovea, a very small spot on the retina that is cluttered with cones. Each of these cones is connected to their own neural pathway. Compared to rods, the color-sensitive cones are relatively less sensitive to light and thus need more time in order to pass the signal along. Foveal vision describes the region of sharp and, above all, conscious image perception. That is why eyesight is usually determined based on foveal vision. Visual acuity and appropriate corrective lenses are also optimized for foveal vision. Objects here are seen as sharp images, similarly to objects seen through a telephoto lens. Foveal vision is used to identify objects. The angle of focused vision can be described as extremely small; only about 2.5 ° are attributable to foveal vision. The area that is perceived as sharp through the fovea corresponds to less than 0.1 % of the entire human field of view. Or, in other words, **at a distance of 2 m, the part of the object that we can perceive sharply is only the size of a pack of cigarettes.**



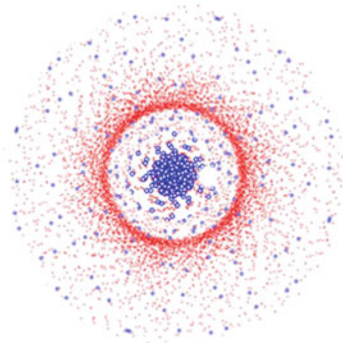


Fig. 3 Schematic structure of the eye [9]; photoreceptor cell density in the eye (qualitative, *blue* rings = cone, *red* dots = stick) distributed over the viewing angle [12]

In this case, the resolution at the edge of the fovea is reduced to less than 50 % of that in the central area.

As mentioned above, focused vision/perception should be regarded as a rather slow system. After directing our gaze onto an object, we need at least 200 ms before the content in this visual field can be identified for the first time [14]. According to Yarbus [15], the typical fixation duration while driving is between 200 and 800 ms. The focus can only be directed towards the next stimulus after this time elapses. Concentrating on multiple objects simultaneously (in parallel) is not possible; the actions can only be performed successively (sequentially). Switching from one to another fixation point (the so-called saccade) involves a realignment of the central line of sight, i.e. a movement of the eyeball (and often also a movement of the head and torso).

In nature, foveal vision is commonly used by hunting animals, which also helps to explain its functionality. Locating the prey is not time-critical: An eagle (the symbol for foveal vision) can, for example, fly for a long time through the air and look for potential prey by "scanning" the ground bit by bit. It is important to spot the prey from a great distance to prevent the prey from detecting the predator too early and fleeing (in fact, an eagle can detect a mouse on the ground whilst flying at an altitude of 1 km). After being identified, the prey remains focused in the foveal vision of the predator right until the point when it is killed. Throughout this process, the environment is not necessarily of interest to the predator. What matters is to spot the target at a great distance, focus on it and not let it out of sight. This approach is often compared to "tunnel vision" or "looking through binoculars".

The so-called "seeing out of the corner of your eye" takes place all around the foveal vision. This process—termed "peripheral vision" in specialist vocabulary—is responsible for ensuring our "overall visual perception". Human beings perceive over 99.9 % of the visual field peripherally. However, since only about 50 % of the data lines (nerves) to the visual center of the brain are allocated to peripheral vision (with the same percentage allocated to foveal vision), we can conclude the following about the functionality of peripheral vision: The information is fed to the brain in a compressed manner. This is done through the linking of several photoreceptors to a single nerve fiber. As a result of this linking, the already light-sensitive cones (which make up most of the peripheral vision) can be exposed to light for even shorter amounts of time. In addition, the linking occurs in such a way that movement stimuli are given preference. The result is very fast but blurred vision, which is particularly sensitive of movement direction and velocity. In nature, peripheral vision is more pronounced in flight animals. For example, an animal that is emblematic for peripheral vision is the hare. As flight animals, hares must be able to quickly detect any predators that approach their danger zone. Visual acuity is not the topmost concern for hares. They only need to detect their hunter, who is usually larger, from a relatively short distance. The large field of view and good overall visual perception give hares a decisive advantage when they try to escape. Hares can achieve perfect spatial orientation and are simultaneously able to perceive objects other than the predator without having to shift their focus towards these objects. This also helps to contextualize the escape tactics of hares.

They zigzag and change direction quickly. The hare escapes from the foveal field of view of the pursuers, who must now spend time to re-orient themselves.

The examples show that the two types of vision serve different purposes. With foveal vision, perception is focused onto one point. In humans, this point is often the object that is actively perceived, or the object on which we concentrate. This is not the case with peripheral vision: Perception in the periphery is normally of a subconscious nature. Both types of vision may be regarded—to the extent that any comparison between the two could be useful—as equally important. Without foveal vision, it would be impossible to identify an object or, for example, read a book. Without peripheral vision, humans would be virtually blind and unable to orient themselves in 3-D space [16]. The properties of foveal and peripheral vision are collated in the Table 1.

Human perception utilizes the advantages of both types of vision. Unless we intentionally scout objects in the far distance, the process usually involves the peripheral detection of objects and then the identification of these objects with foveal vision. One thing to note here is that the pre-selection as to which objects we should focus on with foveal vision is made subconsciously in the periphery. The decisive factors for this pre-selection process comprise the intensity of the stimulus (the more intensive the movement/brightness of the object, the more likely it is for us to focus on it), as well as the expectation and experience of the respective person.

One of the important factors for efficient object recognition in the periphery is "optical flow", which is also known from image processing in advanced driver assistant systems. "Optical flow describes the movement patterns of surfaces and edges in a sequence of images". If objects move differently from the actual optical flow, they are perceived as a stimulus, which usually results in a reflex, in which the eyes (and the body) reorient themselves towards this stimulus [9]. A continuous

Table 1 Properties of foveal and peripheral vision

	Peripheral vision	Foveal vision
Percentage of visual field	> 99.9 %	< 0.1 %
Percentage of optical nerves	~50 %	~50 %
Type of sight	Brightness (black and white)	Color vision
Photosensitivity	High photosensitivity	Low photosensitivity
Visual acuity	Low to very low visual acuity	High visual acuity
Cognitive level	Object detection	Object identification
	Good motion detection	Good object tracking (fixing)
	Random detection in a large visual field	Targeted "scouting" in a very small visual field
	Parallel perception of objects	Fixed perception of one object
	Orientation in space	
Evolutionary purpose	Detection of objects/danger	Fixation onto objects/hunting

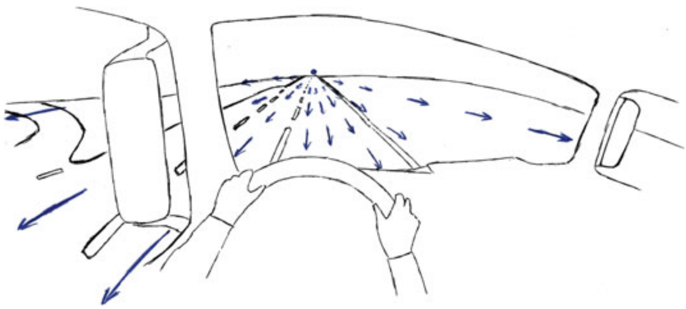


Fig. 4 Optical flow - expansion from the perspective of a truck driver on the country road

(primary) optical flow is important for efficient object recognition. Each local rotation or modification of the optical flow, as happens, for example, in a mirror, means additional effort on the part of the cognitive functions responsible for sight.

The optical flow that arises from a person's own motion has a "far point", at which no relative speed can be recognized. When the observer moves closer to this far point, the optical flow will move towards them (see Fig. 4), and when the observer moves away from this point, the optical flow also moves away. From a physical standpoint, the position of the resting point defines the person's movement direction.

The far point is extremely important for human orientation. Foveal vision is naturally often directed towards the far point (the endpoint in the direction of movement). The optical flow runs evenly towards or away from this point.

Since optical flow is essentially based on the comparison of successively recorded individual frames, a saccade will disrupt the optical flow for a short period. High loads on the visual system (caused by many relevant visual stimuli) may result in diminished visual perception due to the frequent switching of focus. The peripheral region can thus fundamentally be used for recording information, yet the higher/complex the information density, the more peripheral capacity is required in order to account for the selection of the next fixation object; as a result, less to no capacity is allocated for the peripheral recording of information [17]. This condition is called "tunnel vision" and is the "peripheral manifestation of central overload" [7, 18].

3 Human Perception as Direct and Indirect Vision in Commercial Vehicles

In this section, the insights relating to human perception, and, in particular, to foveal and peripheral vision, are used to describe direct and indirect vision in a commercial vehicle. Figure 5 shows the simulated perception of the driver while looking from the vehicle for a duration of 5 fixations (about 1.5–5 s of driving time). It is evident that the driver can only obtain a crisp image of a very miniscule

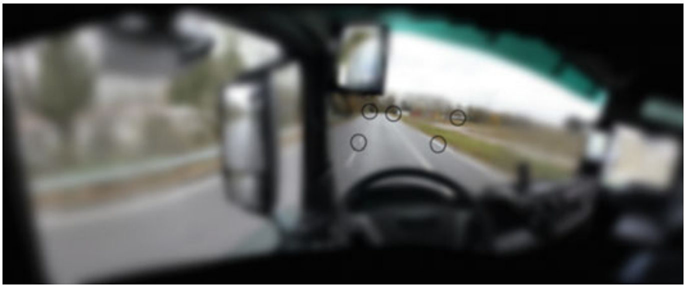


Fig. 5 Simulated human perception while driving (5 fixations) [2]

part of their environment. Thus, human beings create a mental model of their environment from a small number of fixation points, the overall perception of their periphery and based on information derived from past experience.

Of course, this model of the environment accounts not only for the direct field of view but also for indirect vision. In this regard, the peripheral detection of objects in a mirror is very difficult due to the different optical flows (mirrored and interrupted, see Fig. 10) in the main viewing direction and in the various mirrors. Instead of monitoring one global optical flow, the driver now has to keep track of 7 different optical flows simultaneously.

Consequently, the currently used mirror systems entail that the surroundings have to be monitored through foveal vision (also demonstrated in [19] [20]). This, in turn, means that drivers using the current mirror systems must deliberately direct their sight to the respective mirror in order to see the areas to the side and back of the vehicle. This is not an issue for the primary rear view mirror in the truck because it is directed into the distance (to the rear) and thus corresponds perfectly to the evolutionary purpose of foveal vision, namely the targeted localization and perception ("spotting") of distant objects (near the far point).

However, the peripheral detection of objects in the wide-angle mirror is unlikely and even less likely in the close proximity mirror. In wide-angle and close proximity mirrors, the driver's task, namely the reliable, random detection of objects in the vicinity (as described above in the example with the hare), which would ideally be accomplished with peripheral vision, must now be completed with foveal vision, i.e. with a conscious and time-consuming look at the respective mirror (see Fig. 6) [9].

Since it is only possible to focus in one viewing direction—foveal vision cannot work in a parallel fashion—the successive inspection of all individual mirrors requires a certain amount of time. The driver needs, for example, at least 2 s to check the three mirrors on the passenger side [19]. In the meanwhile, at a speed of 30 km/h, the truck has already moved about 17 m forward (by way of comparison: the prescribed length for a class V field of view is currently 4.75 m, which would have been exceeded 3.5 times in this timeframe). Especially in complex traffic situations, drivers are forced to limit themselves to the fields of view that they need

Fig. 6 Foveal look at the mirror: The cyclist is only visible when looking directly at the wide angle mirror (*right picture*). The cyclist remains undetected when looking at the main rear view mirror (*left picture*). [2]

for driving the vehicle (or, in other words, "for accomplishing their objective"). It is up to the driver to decide which mirrors they look at and which mirrors they leave unchecked.

The decision as to which mirror should be looked at is based on the objective by nature, i.e. it is motivated by foveal vision. In studies that tracked the eyes of volunteers in real traffic, the following has been observed and summarized:

The drivers in the tests usually tried to locate the rear axle of their trailer in the mirror (see Fig. 7). This helped them identify if they could pass through narrow stretches and roads (= primary objective of the driver), especially when navigating around turns. If the tail end or axle of the trailer were no longer visible in the main mirror, the drivers either extended their field of view by moving physically toward the front or they looked at the wide-angle mirror [21]. Taking into account the frequencies with which the drivers looked at the different mirrors, it can be stated that the main mirror is used much more often than the wide angle, close proximity and front mirrors.

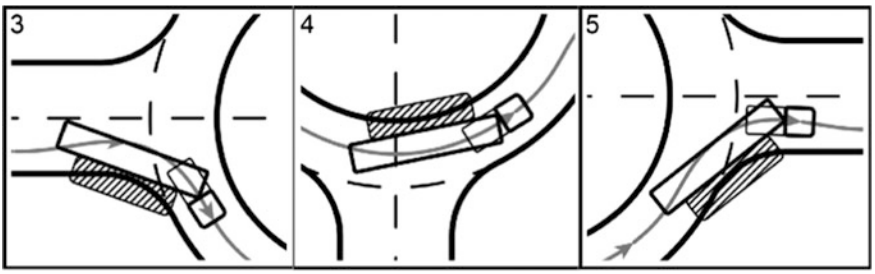


Fig. 7 Viewing targets of drivers passing through a roundabout [21]

The findings from the drive tests once again underscore the foveal motivation of drivers when looking at the mirror. The monitoring of surroundings (field of view classes IV–VI) represents an additional task for the driver (which usually does not coincide with the primary objective) that requires a lot of training simply because it is not a scenario that has been accounted for by evolution.

4 Design of Camera Monitor Systems on Commercial Vehicles

Based on the foregoing observations, it is clear that the mere replacement of each respective mirror with a CMS cannot be effective. Replacing the 6 mirrors with 6 individual frames would largely present drivers with the same challenges as those in the current mirror system. Drivers would again be forced to choose which image they look at and what areas they leave unmonitored. Instead, the aim should be to reduce the number of images displayed while providing the same amount of information to the driver.

Insertion:

Systems that merely replace the individual mirrors have already been presented. A so-called 1:1 replacement for field of view classes II and IV is shown on the left in Fig. 12. *Some of these systems have a feature that tracks the camera of the main mirror replacement based on the edge of the trailer. This enables the driver to monitor the rear axle of the trailer at all times without having to switch to the wide-angle mirror. With regard to comfort, this is clearly advantageous for the driver, but may possibly also have critical consequences. This function causes drivers to pay even more attention to the class II monitor to the detriment of monitoring their surroundings in field of view class IV. Consequently, certain viewing areas—including areas that are particularly important for vulnerable road users—would only rarely be monitored by the driver.*

Based on the results of the drive tests and the findings about human perception, we can formulate the two general **primary aims for** the design of a **mirror replacement** system:

- **The presentation of distant views (i.e. the main mirror area) should be designed in line with the current system of mirrors:**

The view in the main mirror corresponds to the evolutionarily developed foveal purpose, namely targeted scouting into the distance.

- **The presentation of the close-range areas (wide angle and close proximity mirrors) should be designed in a way that ensures the reliable detection of objects.**

From an evolutionarily standpoint, the monitoring of the viewing areas represents a peripheral task, and should be performed using peripheral vision as far as possible.

The two primary aims can only be fulfilled if the number of monitors is reduced. This is achieved by merging the fields of view. A prototype showcasing this technology has been realized as part of a research project between the Institute of

Ergonomics at the “Technische Universität München” and the MAN Truck & Bus AG. This type of visual presentation has already been demonstrated at the Commercial Vehicle IAA 2014 and in various publications [1, 2, 22].

The imaging method can be best explained using the example of a mirror with an aspherical lens, which is known from the passenger car sector. The image from the main mirror area is displayed undistorted on the monitor after applying the required minimum magnification factors. This area is adjoined by an outer compressed area, which roughly represents the missing wide-angle region on the side. The remaining portions of the wide-angle and close-proximity fields of view are also shown as compressed images under these two areas (see Fig. 8). The CMS fulfills the requirements of ISO 16505 according to field of views, imaging and magnification factors [23]. The direct comparison between the mirror system and the CMS is shown in Fig. 9.

The advantages of this view are obvious: The driver is shown a single comprehensive image that contains all of the information that they would otherwise have to put together from three mirrors. This improves the driver’s orientation and the spatial localization of objects. The result is more efficient perception with less time and effort. The driver no longer needs to decide which of the three mirrors they have to look at and can instead obtain all the necessary information at a single glance.

Drivers are more likely to notice objects by using the CMS. Instead of monitoring three different optical flows in the previous mirror system, the entire scene can now be seen as one continuous optical flow on the display (see Fig. 10). The driver can thus focus on their objective using foveal vision (looking into the distance, or detecting the rear axle) while keeping all their close-range surroundings on the sides in their peripheral vision. Objects can be detected much more quickly, more efficiently and, above all, even “randomly”. The “task, which would ideally be accomplished with peripheral vision”, or the monitoring of areas IV and V, can now be accomplished peripherally with the new system. In addition, the CMS allows for

Fig. 8 Combination of the three side mirrors in a single monitor [1]

Fig. 9 The fields of view of all three mirrors are shown in one continuous display

Fig. 10 Optical flow comparison: In the mirror system, the driver must monitor 3 different optical flows; with the CMS, there is a continuous optical flow through all fields of view on the passenger side

the improved spatial recognition of objects outside the main mirror range of the vehicle (see Figs. 11 and 12). The increased detection rate and improved overview in turn also contribute towards improved safety.

Fig. 11 Foveal glance with the CMS: The cyclist is detected at first glance (via peripheral vision) and his position can be determined properly (compared to Fig. 6) [2]

Fig. 12 Comparison of different presentation concepts (*left to right*): 1:1 replacement; 1:1 replacement with camera panning; combined replacement; combined replacement with maneuver view

5 Situation-Adapted Imaging—The "Maneuvering View"

In addition to the above-mentioned default view that is shown to the driver in most driving situations, MAN's CMS also implements a separate maneuvering view. In this view, which can be selected freely by the driver at low speeds, the viewing area on one side is shown on the monitor without field distortion. The maneuver view can be extremely helpful for the driver if, for example, the trailer bends while maneuvering and thereby obscures the main field of view. The maneuver view also provides the driver with a quick undistorted overview of what is happening on the side of the vehicle. The CMS features an automatic switching from default to maneuver view if a pre-defined bending angle between the trailer and the truck is reached. If the driver navigates a tight curve or turns sharply, the inner side of the curve is shown in the maneuver view from the point at which the trailer covers the entire class II field of view. The automated switching process ensures that the driver is always shown an optimized field of view.

Figure 12 shows the different HMI concepts for the design of mirror replacements in commercial vehicles. The 1:1 replacement assumes a display with a format of 8:3, while the 4:3 format of the prototype was used for the combined view. Both displays have the same diagonal measurements. However, the display for the combined replacement has a larger area due to the different format.

In the conventional 1:1 replacement, the trailer obscures the viewing area (as is also the case in the current mirror system). The driver is forced to shift their attention to the wide-angle image area. The cyclist can only be detected after this action is performed.

With a 1:1 replacement using camera tracking for the main mirror area, the rear axle of the trailer and the traffic behind it are clearly visible. The driver no longer needs to look at the wide-angle area in order to determine whether the trailer will make it around the turn. If the driver's view, however, is not directed towards the wide-angle screen, the cyclist will be overlooked completely. In this case, the driver would have to be made aware of the cyclist through additional sensors.

In both 1:1 replacement solutions, an additional mirror or CMS on the passenger side is essential for viewing area V.

In the third image in Fig. 12 (featuring the combined view presented in this paper), the trailer again obscures the entire class II field of view. However, the driver here can immediately detect the cyclist at a glance. Nonetheless, accurate maneuvering around the turn is made difficult by the various distortions in the four areas.

In the maneuver view (right) the entire scene to the side of the vehicle is shown undistorted. Thanks to the undistorted view, the driver can easily maneuver around the turn: The maneuver view displays the bending angle of the trailer and the overall situation in an ideal manner. The cyclist is detected immediately and their position in relation to the truck can be assessed correctly. Thus, the maneuver view is not only useful when maneuvering, but it can also (optionally) assist the driver while driving in the city.

6 Summary

The foregoing shows that there is a huge potential in mirror replacement solutions, provided that these solutions are designed ergonomically, or in line with human characteristics. If the system is optimized for human perception and eyesight, it can not only promote the well-being of the driver but also lead to improved safety on the streets. The combined presentation for field of view classes II, IV and V as shown in this chapter is optimized for foveal and peripheral human perception. The intuitive layout of the image in CMS makes it easy for the driver to analyze objects in relation to the vehicle and to follow these objects beyond the boundaries of the respective viewing areas. The driver is offered a single comprehensive image featuring all the information that they would otherwise have to acquire from three individual mirrors on each side. This facilitates the driver's perception of the areas beside and behind the vehicle because no additional effort is required to integrate the perceived information. Another advantage of this system is that the driver no longer has to make a decision—before even looking at the mirror—as to which fields of view they need to check and in what order they need to check them. All fields of view from one side can be viewed simultaneously, which clearly improves the odds of spotting objects through the driver's peripheral vision.

References

1. Zaindl, A., Zimmermann, A., Dörner K., Kohrs, C.: Camera-Monitor System as Mirror Replacement in Commercial Vehicles, *ATZ,* pp. 12-17, 05 2015
2. Zaindl, A.: Spiegelersatz am Nutzfahrzeug, *Ergonomie Aktuell,* pp. 30–33 (2015)
3. UN/ECE, Regelung Nr. 46 der Wirtschaftskommission der Vereinten Nationen für Europa (UN/ECE)—Einheitliche Bedingungen für die Genehmigung von Einrichtungen für indirekte Sicht und von Kraftfahrzeugen hinsichtlich der Anbringung solcher Einrichtungen *(*2009)
4. Summerskill, S.: The identification of 'Blind Spots' in direct and indirect vision for Category N2 & N3 vehicles using Sigital Human Modelling (2011)
5. Zaindl, A., Zimmermann, A.: Mirror Replacement in Trucks. In: IQPC Automotive Cockpit HMI 2013, Bonn (2013)
6. Neue Lkw-Spiegel: Der tote Winkel ist tot (2004), http://www.spiegel.de/auto/werkstatt/neue-lkw-spiegel-der-tote-winkel-ist-tot-a-323955.html. [Zugriff am 09 Sept 2015]
7. Woyna, L.: Fahrerzentrierter Kriterienkatalog zur Außensichtbeurteilung im PKW. Technische Universität Darmstadt, Darmstadt (2013)
8. Vollrath, M., Krems, J.: Verkehrspsychologie, Stuttgart, Kohlhammer GmbH, pp. 26–40 (2011)
9. Hunziker, H.-W.: Im Auge des Lesers: foveale und periphere Wahrnehmung—vom Buchstabieren zur Lesefreude. Transmedia Stäubli Verlag, Zürich (2006)
10. Gralla, V.: Peripheres Sehen im Sport—Möglichkeiten und Grenzen dargestellt am Beispiel der synchronoptischen Wahrnehmung. Ruhr-Universität Bochum, Bochum (2007)
11. Böhm A., Lux, A.: Augencentrum Dresden (2014), http://www.augencentrum-dresden.de/patienteninfo/Aufbau-des-Auges.php Accessed 01 Aug 2014
12. Funk, G.: Digital Media For Artists Archiv (2002)–(2011) http://www.dma.ufg.ac.at/assets/16457/intern/netzhaut.jpg. [Zugriff am 13 Oct 2015]
13. Lingg, W.: Scheffel Gymnasium Lahr/Schwarzwald (2014), http://www.scheffel.og.bw.schule.de/faecher/science/biologie/sinninhalt5.htmAccessed 07 Aug 2014
14. Werner, H.: Studies on Contour: I. Qualitative Analyses. Am. J. Psychol. **47**(01), 40–64 (1935)
15. Yarbus, A.L.: Eye Movement and Vision. Plenum Press, New York (1967)
16. Leydhecker, W.: Augenheilkunde, pp. 133–135. Springer, Berlin (1990)

17. Cohen, A.S.: Visuelle Orientierung im Strassenverkehr. Bfu, Schweizerische Beratungsstelle für Unfallverhütung, Bern (1998)
18. Cohen, A.S.: Visuelle Informationsaufnahme des motorisierten Verkehrsteilnehmers, Interdisziplinäre Unfallforschung: Informationsaufnahme des Verkehrsteilnehmers, pp. 17–34 (1994)
19. Rassl, R.: Visuelle Informationsaufnahme des motorisierten Verkehrsteilnehmers—Visuelle Informationsaufnahme des motorisierten Verkehrsteilnehmers. Technische Universität München, München (2004)
20. Hudelmaier, J.: Sichtanalyse im Pkw—unter Berücksichtigung von Bewegung und individuellen Körpercharakteristika. Technische Universität München, München (2003)
21. Bothe, A.: Analyse dynamischer Sichtsituationen zur ergonomischen Auslegung von Kamera-Monitor-Systemen (KMS) in schweren Nutzfahrzeugen. Technische Universität Darmstadt, Darmstadt (2014)
22. Bengler, K., Götze, M., Pfannmüller, L., Zaindl, A.: To see or not to see- innovative display technologies as enablers for ergonomic cockpit Concepts. In: Electronic Displays Conference, Nürnberg (2015)
23. ISO 16505:2015, Road vehicles—ergonomic and performance aspects of camera monitor systems—requirements and test procedures, International Organization for Standardization (2015)
24. Schulz, R.: Blickverhalten und Orientierung von Kraftfahrern auf Landstraßen. Technischen Universität Dresden, Dresden (2012)
25. Remlinger, W.M.: Analyse von Sichteinschränkungen im Fahrzeug (Dissertation). Technische Universität München, München (2013)

Ergonomic Design of Camera-Monitor Systems in Heavy Commercial Vehicles

Alexander Georg Bothe

Abstract To fulfill the transportation task the driver has to steer the commercial vehicle through a challenging transport infrastructure. The driver-related vision issues, particularly the indirect vision to monitor vehicle behavior during dynamic vision situations are of particular importance. The main requirements of the haulage contractor, when making the decision to buy a new vehicle, are a cost-effective, fast and flexible transportation of goods. When designing commercial vehicles all requirements are to be considered. Caused by the potential for fuel savings and the technology-related benefits for the truck drivers, it is of interest to replace conventional mirror systems of a commercial vehicle with camera-monitor systems (CMS). The electronic vision system is directly involved in the dynamic viewing interaction between driver, vehicle and environment and therefore needs to be designed according to ergonomic methods. The findings of the analysis of dynamic vision situations are used for the ergonomic design of a camera-monitor system (CMS). The viewing angle-dependent image quality of the monitors is taken into consideration for the display adjustment depending on the dynamic eye point positions. The automated adaptation of the shown dynamic fields of vision is implemented through a truck-trailer angle dependent panning characteristic of the CMS. Thus, the fuel saving advanced vision system is adapted to the ergonomic needs of truck drivers.

Keywords Ergonomic design of Camera-Monitor systems · Heavy commercial vehicles · Transportation task · CO_2-Reduction · Dynamic vision interaction · Vision situations · Vision task · Vision restriction · Monitor alignment · Panning characteristic

List of Abbreviations

AOI	Area of Interest
CMS	Camera-Monitor System
CO_2	Carbon Dioxide

A. Georg Bothe (✉)
Institute of Ergonomics, Darmstadt University of Technology, Darmstadt, Germany

A. Terzis (ed.), *Handbook of Camera Monitor Systems*,
Augmented Vision and Reality 5, DOI 10.1007/978-3-319-29611-1_11

dyn	dynamic
ECE	Economic Commission for Europe
FOV	Field of View
GIP	Gaze Intersection Point
hor	horizontal
ISO	International Organization for Standardization
neg	negative (−)
ORP	Ocular Reference Point
pos	positive (+)
SgRP	Seating Reference Point
stat	static
ver	vertical
VR	Virtual Reality

1 Introduction

1.1 Motivation

The demands on the driver during the transportation task in commercial vehicles differ significantly from the driving task in passenger cars. To fulfill the transportation task the driver has to steer the commercial vehicle through a challenging transport infrastructure. The main requirements of the haulage contractor, when making the decision to buy a new vehicle, are the cost-effective, fast and flexible transportation of goods. When designing commercial vehicles both requirements are to be considered.

For the variety of different transportation tasks optimized vehicle types are designed for the individual use cases. In Fig. 1 heavy long-distance transport vehicles are shown in a typical driving situation on the highway. During the long-distance drive the driver assesses lane keeping and distance estimation regarding vehicles ahead. When maneuvering, the monitoring of the direct vehicle environment is particularly important. Compared to passenger cars the size and complexity of modern highway trailer trucks are particularly noteworthy. The driver-related vision issues from the high eye point position is therefore of great importance in commercial vehicles.

For the commercial vehicle industry CO_2 reduction targets are defined through regulations of the legislature and commitments by the vehicle manufacturer. A reduction of CO_2 emissions and the associated reduction in fuel consumption is also an important factor in the buying decision of the haulage contractor. Cameras outside the vehicle, an electronic signal processing system and monitors inside the cabin, can replace the conventional mirror system. Thus, the cross-sectional area and the fuel consumption of commercial vehicles are significantly reduced by using

Fig. 1 Heavy commercial vehicles in long-distance haulage [19]

camera-monitor systems (CMS) for indirect vision. The electronic vision system is integrated directly into the dynamic vision interaction between driver, vehicle and environment and must therefore be designed according to ergonomic methods.

With the technology change necessary to implement a CMS a change in the information flow within the vision interaction of the driver with the target via the respective vision system is induced. For mirror systems, the optical signal transmission between target and eye point is linked directly via the mirror. For camera-monitor systems (CMS), there is no direct optical link between target and eye point. An initiated head movement of the driver that changes its eye point position results in a shift in the visible area shown in the mirror. In case of a simple CMS, the same head movement has no influence on the display field of indirect vision.

For an ergonomic design of CMS the dynamic use of conventional mirror systems must be considered in more detail. By analyzing the vision interaction during dynamic vision situations vision areas used by the driver and the used eye point positions could be determined. The dynamic vision data enables a design of camera-monitor systems (CMS), which takes into account the requirements of the truck driver to fulfill the transportation task.

1.1.1 Transportation Task of Commercial Vehicles

As a base for describing the transportation task within the commercial vehicle logistics the driving task of [65, p.18] is used: "The driving task consists primarily of the people steered movement with a land-based motor vehicle from A to B, taking into account that no other objects are affected except the road […] Additional, individual aspects of the driving task are e.g. the fact that the driver wants to be particularly economical, sporty, comfortable etc. on the go." In European long-distance commercial vehicle environment, the transportation task from A to B is planned by the haulage contractor and not by the driver [43, p. 187]. Main requirement of the haulage contractor is the cost-effective, fast and flexible delivery of goods [35, p. 1]. The haulage contractor thus specifies that the professional driver has to use the vehicle as economically as possible. Comfort

Fig. 2 Reference semitrailer-tractor combination [19]

preferences or temporary power requirements are therefore not in the first place when addressing the transportation task.

For the wide variety of transportation tasks vehicle types optimized considering the use case are developed. It can be distinguished between long-distance haulage, distribution and site vehicles [35, p. 1], [79]. The long-distance traffic environment is divided by the vehicle combination between single vehicles, drawbar and semitrailer-tractor combinations [12, p. 111], [13, p. 31]. The reference semitrailer-tractor combination illustrated in Fig. 2 represents the most important long-distance transport vehicle in the European market [68, p. 19]. For urban driving or maneuvering direct direction changes must be carried out on limited infrastructure area with those vehicles. Semitrailer-tractor and drawbar combinations are designed on the basis of this requirement and their total length as a combination of the drawing vehicle and the trailer. The bending of the trailer during direction change increases the complexity of the vehicle dynamics and complicates the allocation of the individual vehicle elements in monitoring the lane keeping by the driver.

Since the 80s, the systematic cost optimization of manufacturing facilities leads to a reduction of internal storage. Through a close relationship between producer and logistics [43, p. 245], the necessary storage capacity is distributed on the "just-in-time" delivering trucks or the nearby located logistics centers. Through this close link between production and logistics, a high economic damage may occur in case of failure of one link of the supply chain. The fleet and time management, as well as the necessary telematics components [42, p. 945] are gaining in importance for the haulage contractors [30, p. 5]. The driver represents in complex traffic with still increasing traffic density [50, p. 53] and strictly regulated driving hours [85 A 6] a high cost source and a potential default risk. However, he is also one of the largest remaining influencing factors for reducing fuel consumption [80, p. 1].

The importance of the transportation task in the concept design of a commercial vehicle is particularly apparent when the geometric dimensions of the reference

Fig. 3 Comparison of direct vision from a midsize sedan and the semitrailer-tractor combination [19]

tractor-trailer are considered (Figs. 2 and 3). The maximum permitted length, width and height dimensions [89, Sect. 13] are fully implemented [14, p. 60]. The cargo compartment of the trailer represents the greatest volume. The tractor is designed so that it provides a very short portion of the total combination [28, p. 26]. Therefore, the dimensional concept of the cabin is optimized to a low length component. The concept related eye point position of the truck driver is determined by this prioritization [17, p. 43]. A variation of the eye point positions caused by the body dimensions of the drivers mainly takes place in the vertical vehicle direction.

In Fig. 3 the static eye point positions and the driver's vision cones of a midsize sedan and the reference semitrailer-tractor combination are shown. The vehicles are placed on the road plane and normalized in the longitudinal direction on the eye point position. The great distance between the two eye point positions in vertical direction and the different window opening areas are obvious [17, p. 44]. The truck driver has in contrast to car drivers, caused by the vision obstruction created by the cargo volume, no possibility of direct vision to the rear. The area directly in front and to the side next to the driver's cab can't be seen directly, due to the high attachment height [18, p. 42], [71, p. 9]. These vision restrictions have to be compensated by the truck drivers through a dynamic vision behavior using the mirrors.

The truck has due to its mass of 40 t [89, A1], the limitation of the maximum available drive speed of 90 km/h [13, p. 74] and the need for economic fuel use [80, p. 1] limited acceleration ability to perform. The driver using the optical distance estimation between his vehicle, other road users and the lane markings must therefore move the truck with foresight. Due to the vehicle dimensions and the complexity of the combination, the lane keeping of the trailer must be checked when performing lateral dynamic driving maneuvers [69, p. 14]. To fulfill the transportation task, the larger and more complex commercial vehicle has to be steered by the driver through the same transport infrastructure, as the relatively small, agile and compact passenger car. The truck doesn't provide direct vision to the rear, so the driver has to monitor lane keeping and distance estimation of approaching traffic via vision systems for indirect vision [71, p. 9]. Thus, the indirect vision in commercial vehicles has a higher priority compared to passenger cars and requires special attention when designing those vehicles.

1.1.2 CO_2 Reduction Targets as a Trailblazer for New Technologies

The European Commission has presented a roadmap for reducing CO_2 emissions in 2011 [88, p. 4]. Accordingly, by 2050 the emission of greenhouse gas has to be reduced by 80 % (base year 1990). The increase in efficiency of long-haulage traffic passes a key role to reach the reduction targets [87, p. 7]; [88, p. 7]. In 2008, a commitment by the commercial vehicle manufacturer to reduce CO_2 emissions has been approved by the European Automobile Manufacturers' Association (ACEA) [59, p. 1]. This commitment includes, by 2020, an average reduction in fuel consumption by 20 % per ton kilometer (reference year 2005). Since the real development of the fleet consumption does not meet those targets [72, p. 188], legal limits on CO_2 emissions of heavy commercial vehicles are to be worked out in the four key markets worldwide [26, p. 393]. These regulations, with non-compliance punishable by fines by the automotive industry [26, p. 390] should enter into force in the coming years.

According to Breuer and Kopp [14, p. 60] a long-distance truck drives an average of 150,000 km per year and consumes an average of 49,500 liters of diesel. The price of diesel for large German freight forwarders was in the last three years on average at € 1.34 [22, p. 28]. The freight forwarder has to pay annually € 66,309 of fuel costs for an average truck and could save € 663 per year for a 1 % reduction in fuel consumption. Through this business context it can be estimated how much a measure of known CO_2 savings potential should cost in order to sell it profitably to the freight forwarder.

Vehicle manufacturers have the interest to achieve the predetermined CO_2 reduction targets through appropriate measures. The necessary CO_2 reduction can be realized by reducing the driving resistance. As the proportion of air resistance of the total driving resistance at high vehicle speed is the greatest, aerodynamic optimization measures are most effective for long-distance transport vehicles [44, p. 653].

Aerodynamic measures for optimizing air resistance and their reviewed potentials are described in the literature [21, p. 33], [28, p. 23], [44, p. 675]. A substitution of the main and wide-angle mirror with a camera-monitor system (CMS) is an outstanding measure. The cross-sectional area of the vehicle is reduced and the drag coefficient can be optimized. In addition, wind noise and pollution of the side windows can be significantly reduced by the elimination of the outside mirrors [27, p. 40], [33, p. 536], [44, p.720]. For concept related direct flow of the mirror areas, the proven fuel saving potential of a CMS is estimated at 2.9 % [21, p. 59].

Assuming a real reduction in fuel consumption through a CMS on a production vehicle in customer use of 2 %, the freight forwarder can save with such a system € 1,326 of fuel costs each year. In the current shell truck study an average holding period of 4.4 years is reported for tractors in Germany [68, p. 30]. This leads to a cost saving of € 5,834 for the first customers over the holding period. The freight forwarder will invest a portion of his savings for this measure and thus defines the selling price of the CO_2-reduction measure.

The CO_2 reduction targets can support, in combination with the fuel saving potential, the technology change from conventional mirror systems to new camera-monitor systems (CMS). Through the proven savings it is possible to integrate this vision enhancing system in long-distance transport vehicles. The electronic vision system is directly involved in the dynamic viewing interaction between driver, vehicle and environment and therefore needs to be designed according to ergonomic methods.

1.1.3 Establishment of Camera-Monitor Systems

Cameras outside the vehicle (see No. 20 in Fig. 4; [10, p. 5]), an electronic signal processing system and monitors inside the cabin, can replace the conventional mirror system [84]. Due to the regulations for vision systems for indirect vision it is approved to replace only front and ramp mirrors (Class V and VI according to [84, par. 15.2.4.2] by CMS (see [84, par. 15.2.1.1.2]. However, with the replacement of the main and wide-angle mirrors (Class II and IV) bigger fuel reductions can be achieved [21, p. 59].

The current mirror systems do not only impede the flow around the vehicle but also block the direct view from the vehicle (Fig. 5; [27, p. 41], [71, p. 10]). For the homologation of commercial vehicles in Germany the “Guideline for the vision of motor vehicles” [82, Sect. 35b] has to be fulfilled. This regulation describes the

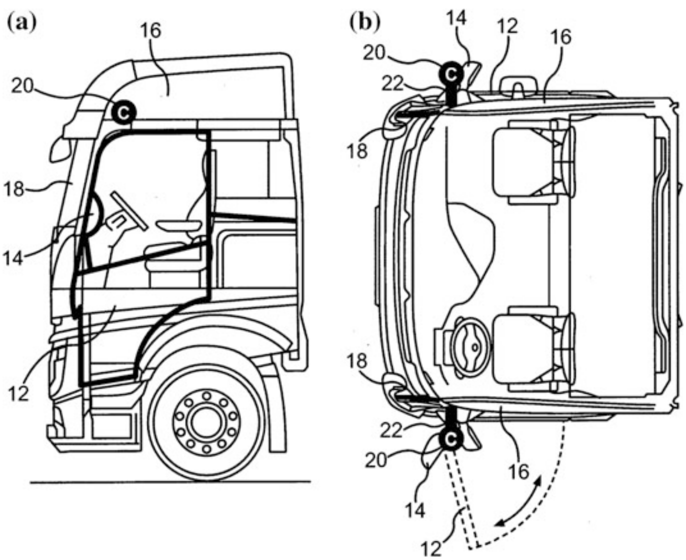


Fig. 4 Components of a camera-monitor system (CMS) for commercial vehiclese [10, p. 5]

Fig. 5 Obstruction of the direct vision through the driver's-side main and wide-angle mirrors on a long-distance transport vehicle [19]

maximum vision angle of obstruction, but mirror devices are excluded in the calculation. The maximum size of mirror systems is therefore not regulated by the legislature, which leads to a continuous increase of the mirror vision obstruction effect [5, p. 5]. When dimensioning vehicle mirrors a conflict exists between the best possible indirect vision and the lowest possible obstruction of direct vision.

In order to satisfy customer expectations regarding the field of vision the necessary mirror surfaces of current mirror systems go far beyond the legally required fields of view [[84, par. 15.2.4.2], [4, p. 1638], [27, Appendix D]). In addition, the aerodynamic optimization of the main and wide-angle mirrors [28, p. 30] leads to large-scale view blocking trim parts [17, p. 97] and a disadvantageous mounting position of the mirrors. Through the establishment of camera-monitor systems (CMS), the described disadvantages can be significantly reduced.

Facing the potential of reducing CO_2 emissions, the reduction of wind noise, enhancing the direct vision and the integration of innovative driver assistance systems, the integration of requirements regarding camera-monitor systems (CMS) into the ECE-R46 [84] has been initiated. For the definition of fundamental technical and ergonomic requirements, a working group that is administered by the International Organization for Standardization (ISO/TC 22/SC 17/WG 2) was established. This working group, including international representatives from government organizations, technical testing services, automobile associations, vehicle manufacturers and system suppliers has the task to define the minimum safety, ergonomics and performance requirements and the associated test methods for a CMS to replace the mirror Class I to IV [84, par. 15.2.4.2] and to document them in the ISO standard 16505 [90, 34, p. 3]). The ISO 16505 is used as the basic technical document amending ECE-R46. By this procedure camera-monitor systems (CMS) are ready for homologation in the foreseeable future and the development is initiated for vehicle manufacturers and system suppliers.

The relevance of the development of electronic mirror replacement systems is demonstrated in the large number of current notifications of patent claims and published patent applications [4, 8, 10, 46–48]. Since the fundamental technical requirements are defined by ISO 16505 and ECE R46 the same for all vehicle manufacturers, the ergonomic driver-related aspects will play an important role for the future market penetration of CMS especially in the commercial vehicle environment. Thus, a recent published patent application [8, p. 1] describes the functionality of controlling the field of vision depending on vehicle parameters. There are suggested control parameters for the automated manipulation of the displayed vision area such as the bend angle between tractor and trailer, the vehicle speed and the steering wheel angle. The active control is necessary because with a camera-monitor system (CMS), in contrast to the mirror system, there is no direct optical link between vision target and eye point. For ergonomic design of CMS initially the dynamic use of conventional mirror systems must be considered in more detail. From the dynamic driver usage the parameters for panning the shown vision areas can be derived.

1.1.4 Dynamic Vision Interaction with Conventional Mirror Systems

Depending on the individual phase of a complex vision situation, such as the roundabout passage illustrated in Fig. 6, different areas of the vehicle surroundings are of particular importance for truck drivers at this time. Those vision areas are controlled by the driver via mirror systems [9, p. 298].

Vision phase 1 is similar to the behavior during straight driving on interurban roads. The driver checks at a relatively high vehicle speed by short monitoring gazes into both main mirrors the straight running of its trailer and its position within the lane. Vision phase 2 represents the driving up to the stop line and the switch to the slow phases of the roundabout passage. The environmental situation with other right of way road users is monitored by means of direct vision. Before entering the roundabout immediate environment of commercial vehicles is checked via the close proximity mirrors. Vision phase 3 is the first of the three main phases of the roundabout passage. It describes a change of direction to the right and is therefore comparable to the vision situation of turning to the right. At low speed, a negative bend angle between tractor and semi-trailer is generated. The driver checked at that time the lane keeping on the right side of the semi-trailer. On the driver's side, the vision onto the relevant areas of the wheels and the lane markings is blocked by the swing-front edge of the trailer.

Vision phase 4 is characterized by a change of driving direction to the left. Depending on the roundabout diameter, this phase can be compared to a left turn or cornering to the left. The result is a positive truck-trailer angle. The relevant vision area is changing from the right to the left side. Vision phase 5 includes a change of direction to the right and the related switch of the vision area to the right side of the vehicle. As in the previous two main phases in phase 5 the relevant field of vision is directly associated with the highly bent semi-trailer. Vision phase 6 describes the transition from the roundabout passage to driving straight ahead. To observe the trailer and the traffic from behind with increasing vehicle speed, the field of vision is

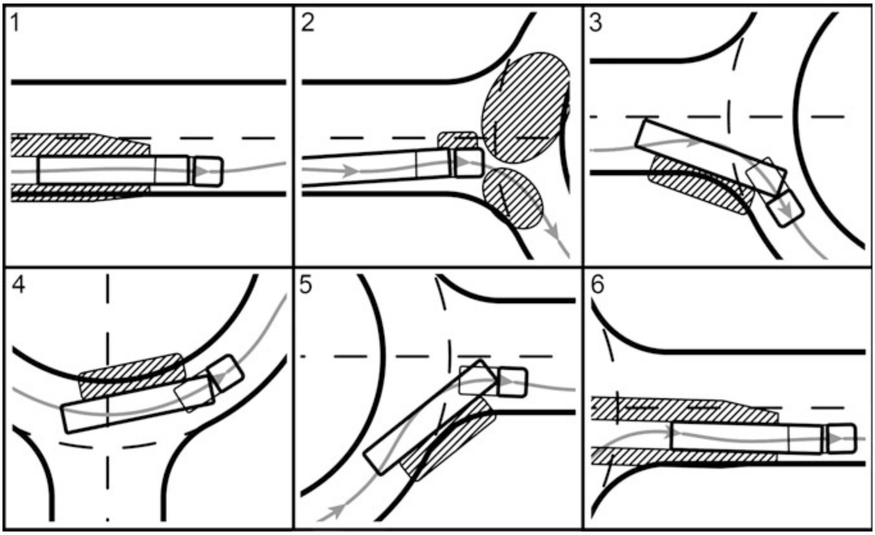


Fig. 6 Dynamic driver used vision areas (*shaded areas*) on the example of the roundabout passage [9, p. 298]

shifted from the near field backwards into the vehicle longitudinal direction and distributes on both sides of the trailer. In summary, for the roundabout passage as a representative of a complex vision situation, a speed dependent panning of the vision area in the vehicle longitudinal direction can be observed. The panning in vehicle transverse direction depends on the bend angle shift of the trailer.

The static fields of vision of a Class II mirror described in the ECE-R46 [84, par. 15.2.4.2] covers only the relevant vision areas of phases 1 and 6 (straight ahead driving). To see the relevant fields of vision of phases with large truck-trailer angles, the driver must adjust its eye points relative to the mirror [4, p. 1638], [27, p. 53]. In Fig. 7 the described dependence between the eye point position, the mirror surface and the shown field of vision is visualized using a CAD simulation. In the left image section, the driving position of the RAMSIS manikin including eye point position and gaze orientation is shown. In the right image section, the rendering shows the driver's view from the eye point in the mirror. In order to evaluate the visual field represented and for allocating to the vehicle the trailer is shown in purple, an overtaking vehicle in green, the field of vision Class II in orange, the field of vision Class IV in yellow and the sky in blue.

The interaction between the driver and the mirror for an intended panning of the shown viewing area is based on a skilled driver's mental model of the visual system and will be applied due to the larger effect mainly in the main mirrors (Class II) and not in the wide-angle mirrors (Class IV). The model includes the positioned eye

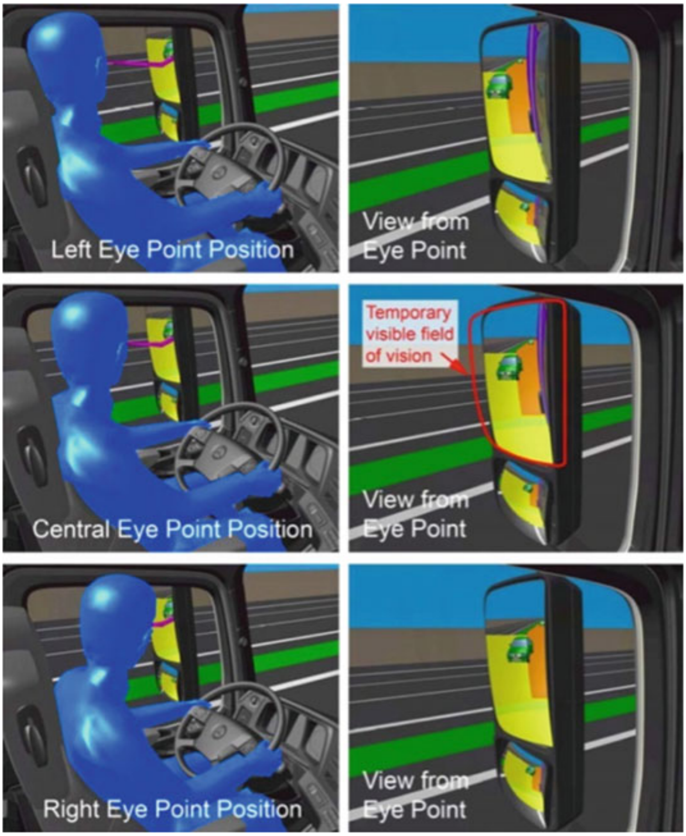


Fig. 7 Simulation of driver-initiated dynamic adjustment of viewing areas in conventional mirror systems [4, p. 1638]

point, the aligned mirror surface and the detected vision target [4, p. 1650]. Since the driver is aware of the optical link between the model components, he may intuitively vary the manipulation variables during a driving situation. The driver is able to see the required field of view with the assistance of specific head movements [4, p. 1638], [18, p. 30], [78, p. 1] within the system-related limits. Within a camera-monitor system (CMS), the optical linking between eye point position and vision target point is interrupted. A head movement of the driver has no influence on the system behavior. In order to provide the drivers required vision areas during the driving task even with a CMS, the situational vision needs can be recorded in the use of mirror systems and reproduced on the monitors of a CMS vehicle parameters dependently [8, p. 1]. The knowledge about the dynamic vision interaction necessary for ergonomic design of a CMS has to be researched.

1.2 Scope of the Study Content

Within this Chapter, the fundamentals for the ergonomic design of camera-monitor systems (CMS) in heavy commercial vehicles are considered. Under heavy commercial vehicles heavy tractor-trailer and drawbar-combinations for the European long-haulage traffic are to be understood [13, p. 31], [79, p. 198]. As shown in Sect. 1.1.2, aerodynamically reasonable measures such as the CMS are only of interest in the long-distance traffic [44, p. 653]. The delimitation of the train combinations is done, as in the European long-distance transport almost exclusively tractor-trailer and drawbar-combinations with more than 12 tons are used [68, p. 19].

The global truck market is divided into several local submarkets, such as Europe, the USA, South America, China, Japan, etc. This situation is mainly owed to the regulations of those submarkets [79, p. 204] such as different length and weight limits between the EU and the USA. The regulations of indirect vision also differ for these target markets [84, 86]. The local regulations lead to different vehicle concepts and combinations, thereby delimitation of this study to the European market is necessary.

Under camera-monitor systems (CMS) only new systems to replace the mirror Classes II and IV [84, par. 15.2.4.2] and not existing systems for replacement of the front mirror (Class VI; illustrated in Fig. 8), are meant [90, p. 1].

Camera-monitor systems (CMS) according to [90, p. 1], [84, par. 2.1.2 & 6.2.2] are vision systems of indirect vision. Therefore, the study scope is limited to the visual interaction of indirect vision. Subjects of direct vision from the vehicle are only treated in order to display the necessary relationships. Technical specifications and methods of measurement for CMS are described in ISO 16505 [90, p. 1] and are therefore excluded from this study. Rather, the outstanding topics of the ergonomic design of camera-monitor systems (CMS) are treated. The orientation of the CMS according to the dynamic eye point positions of all relevant driver types and the definition of the characteristic curve vehicle parameters depending visualization of dynamic vision areas in the CMS is the core application of this study.

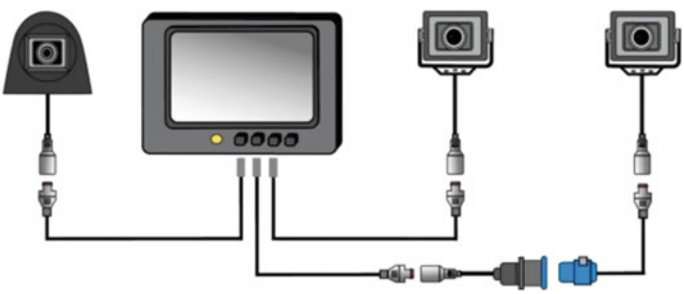


Fig. 8 Camera-monitor system for replacement of the front mirror and for showing the rear view camera (assistance function) for tractor-trailer and drawbar-combinations [51, p. 8]

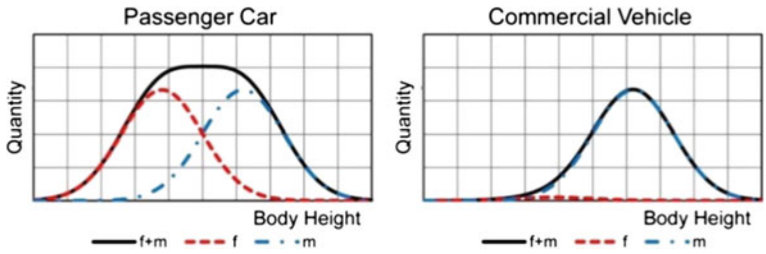


Fig. 9 Influence of the gender distribution of car and truck drivers on the total body height distribution [7, p. 442]

The group of truck drivers differs from the whole population because of characteristics such as the gender distribution of the occupational group [40, p. 1]. Those aspects are relevant for ergonomic design. In contrast to passenger cars, in which an equal weight distribution between female and male customers is recognized the influence of female truck drivers on the entire body height distribution is negligible. Therefore, in this study, the body height is exclusively based on surveys of male populations (Fig. 9; [7, p. 442], [40, p. 1]).

In order to perform the ergonomic design of camera-monitor systems (CMS), data of the dynamic interaction is collected during real vision situations. This data is analyzed application-oriented. The results of these studies are to be used for ergonomic design of CMS to replace conventional mirror systems in typical long-haulage heavy commercial vehicles. Before an application of the findings on the interpretation of other vehicle systems or in other vehicle concepts, transferability should be critically examined.

2 Fundamentals

2.1 Ergonomic Aspects of the Vision from the Vehicle

For ergonomic design of new system components, these have to be adapted to the abilities and skills of people [62, p. 969]. Therefore, this section shows the state of research regarding information acquisition and processing based on established model approaches. To fulfill the task of driving the visual sensory channel is by far the most comprehensive [1, p. 6], [76, p. 1]. The design and study relevant characteristics of the human eye are therefore summarized briefly. This is followed by consideration of the vision task and the related vision restrictions from commercial vehicles.

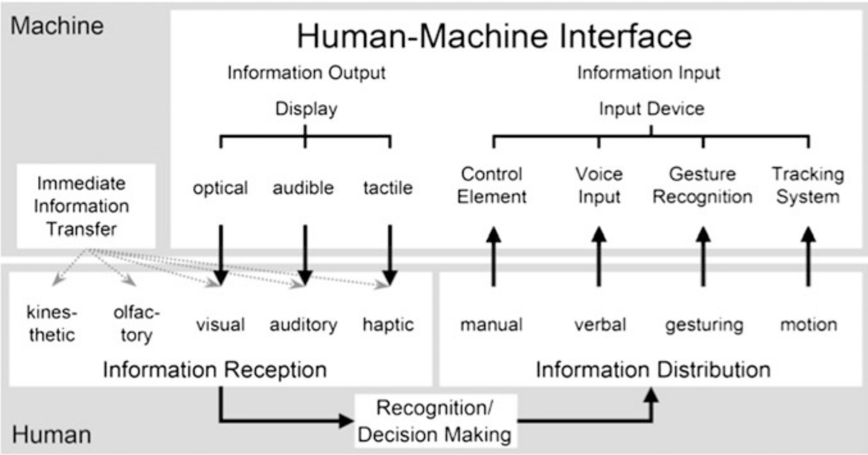


Fig. 10 Information transfer in human-machine interfaces [62, p. 971]

2.1.1 Information Acquisition and Processing

The camera-monitor system (CMS) is in addition to the conventional control and display components [3, p. 38], [61, p. 29] considered as a new element of the man-machine interface of a commercial vehicle. The individual steps (stages of information processing) and sensory modalities (channels of human perception) for transmitting information within a man-machine system are shown in Fig. 10 (see [62, p. 971]).

To describe the human information processing sequential [49, p. 153], [56, p. 258] resource-based [74, p. 10] and hybrid [1, p. 4] modeling approaches are used [62, p. 286]. All these established information processing models have in common that the received signal is converted through the sense organs of man (stimulus) following an internal processing (cognition) in a reaction (response) [1, p. 4]. The stages of processing information can thus be divided into the information recording (perception), the information processing (cognition) and the information output (motor function) [1, p. 5], [62, p. 969]. Since the available information processing resources are limited not all the information but only specifically selected signals can be supplied to the conscious information processing [1, p. 5], [39, p. 402]. With the information output, information processing based actions are realized. In the case of the driving task, for example, targeted movements of the hand-arm or torso-head systems [1, p. 8]. By the described information processing an interaction within the human-machine system takes place.

The human being receives optically information shown in the camera-monitor system (CMS) during a visual situation in the vehicle via the visual sensory channel. Therefore, the visual information is considered as an input variable of the driver-vehicle-environment control loop [58, p. 47]. Remlinger describes the primary driving task using the “three-level hierarchy” according to Donges [23, p. 183], [24, p. 16], [58, p. 47], [65, p. 5]. Within the driving task, information from the environment, from the vehicle and from the reaction of the entire system on the driver's action affect the driver.

In addition to the visual stimuli auditory, tactile, kinesthetic and olfactory signals effect on the driver [15, p. 103], [62, p. 313], [65, p. 4], [76, p. 1]. 80–90 % of the information necessary to meet the driving task is received visually [60, p. 150], [1, p. 6], [31, p. 230], [58, p. 21], 60, p. 29], [76, p. 1]. The eye is, according to Cohen [16, p. 3] the only sense organ that can receive targeted information from distances beyond the vehicle stopping distance. By this characteristic, the prominent position of the visual information received during driving situations can be founded.

2.1.2 Characteristics of the Human Eye

The human eye and its characteristics are of particular importance to fulfill the driving task. Light enters through the pupil, an opening in the iris, into the inside of the eye and meets after lens and vitreous body are passed, on the light-sensitive retina [45, p. 1], [62, p. 317]. The Iris has thereby the function of limiting the incident light quantity and thus adapts the eye to different ambient brightness. Vision objects at different distances from the eye are clearly focused on the retina by changing the lens curvature. This adjustment of the lens focal length is called accommodation. The accommodation capacity decreases with age and results during frequent changes of vision distance in fatigue symptoms [62, p. 318].

If the eye is aligned over the fixation axis onto a possibly moving vision object, it is called a fixation. The eye is related to the view object in relative stability [54, p. 3]. To maintain the sharpness micro movements of the eye are performed [77, p. 107]. The change to a new vision object occurs with a fast ballistic movement (saccade). These movements take place at a rate of up to 900 °/s on average three times per second [54, p. 3], [55, p. 8]. The duration of a saccadic eye movement can be estimated at 20–80 ms [45, p. 8]. During such switch in glance no visual information is processed [36, p. 120]. The fixation duration during the driving task is on average 300–400 ms [41, p. 53], [65, p. 84] and can thus be clearly differentiated from the saccades.

Fixations can be separated on their duration and compared with each other [55, p. 8]. Via the orientation of the fixation axes of the left and right eye the distance of the vision targets can be determined (see Fig. 11) and the necessary accommodation can be derived. Due to the intersection of the fixation axis with a known area of interest (AOI) the exact gaze intersection point (GIP) can be calculated. The peripheral vision is not suitable to describe the directional information received [65, p. 5]. Because of the reduced visual acuity and color perception objects are only detected in the peripheral area [45, 63, p. 10]. In order to derive relevant information concerning the dynamic vision interaction to fulfill the driving task, therefore it is suitable to record dynamic eye point positions and gaze orientations during dynamic vision target fixation.

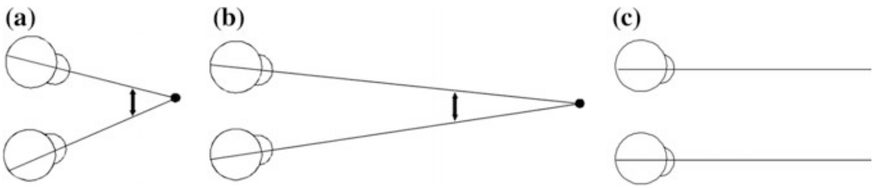


Fig. 11 Fixation to **a** close, **b** medium and **c** far-away vision targets [25]

2.1.3 Vision Task and Vision Restriction for Commercial Vehicles

The vision task in the commercial vehicle is derived from the transportation task described in Sect. 1.1.1. Other road users, transport infrastructure and the behavior of the own vehicle are monitored within the driving task [23, p. 183], [65, p. 18] from the commercial vehicle in order to derive driver reactions [1, p. 4]. The concept related eye point position and the resulting vision restriction have to be considered.

In Fig. 12 the interaction between direct (1) and indirect vision (3–6) from a tractor (7) is shown [90, p. 93]. The vision areas 3–6 represent the legally required fields of vision [84, par. 15.2.4.2]. However, the vehicle manufacturers cover larger vision areas of indirect vision to fulfill the customer expectations [27, Appendix D]. Despite this field coverage large areas that can be seen neither via direct nor via indirect vision occur nearby the cab. These blind spots (2) [17, p. 91], [71, p. 20],

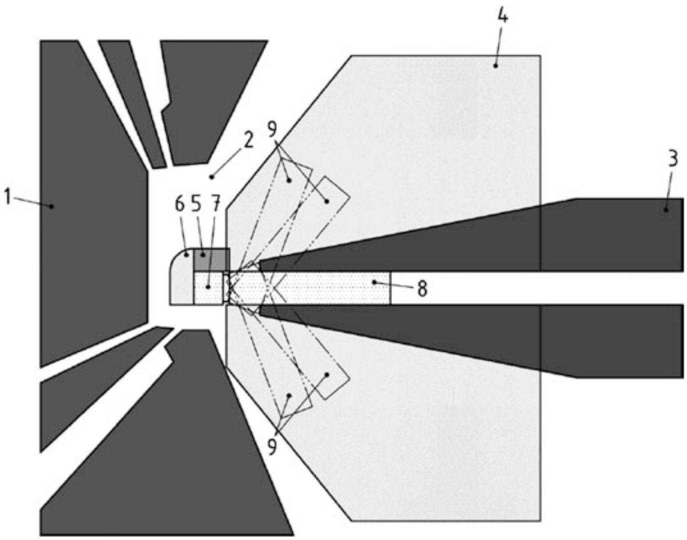


Fig. 12 Interaction of direct and indirect vision in commercial vehicles: viewing areas on ground level [90, p. 93]

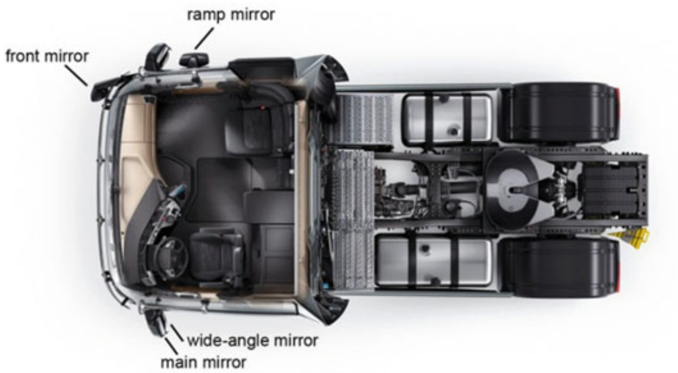


Fig. 13 Mirror assembly mounted on a tractor unit [19]

which are caused by vision restrictions, are considered in more detail below. The shown commercial vehicle consists of the tractor and the trailer (8), which is aligned at a bend angle relative to the drawing vehicle depending on the driving situation. Possible semitrailer positions (9) with large bend angles that typically occur when turning or maneuvering [9, p. 298], [78, p. 2] are also shown [90, p. 93] and must be considered in the ergonomic design of indirect vision.

The available visual information within the vision task is already reduced by the limitations in the information receiving by the human eye and the limits of driver-initiated compensation movements (see [37, p. 21]). Also, the vehicle provides, due to his vision restricting effects an information filter [75, p. 1]. Depending on the design of a semitrailer-tractor combination for long-haulage traffic, the direct vision is very limited (Sect. 1.1.1).

The legally required fields of vision (3 to 6) [84, par. 15.2.4.2] are covered by means of the six vehicle-mounted mirrors (Fig. 13). These are the driver and passenger side main mirrors (Class II) and wide-angle mirrors (Class IV), as well as the ramp mirror (Class V) and the front mirror (Class VI). Compared to the passenger cars, the very high number of necessary mirrors underlines the weight of indirect vision of commercial vehicles. To monitor the relevant vehicle environment the driver must view multiple separately arranged mirrors and evaluate them by using a holistic mental model [32, p. 641] of the vehicle-environment system.

Due to the optical link between eye point position, mirror surface and vision target [9, p. 298], [18, p. 30] the mirrors shall be attached to the vehicle in a way that they are visible from the driving position and display the relevant vision areas to the driver. This is associated with a further reduction of the available areas of body openings usable for direct vision. Thus, the mirrors lead to deterioration in direct vision ([18, p. 39]; Fig. 12). [20, p. 42] can identify the exterior mirrors of a semitrailer tractor as the most important vision masking parts of the vehicle by a subject survey (Fig. 14). Within the study truck drivers were asked for the body parts or attachments, blocking the view from the vehicle cab.

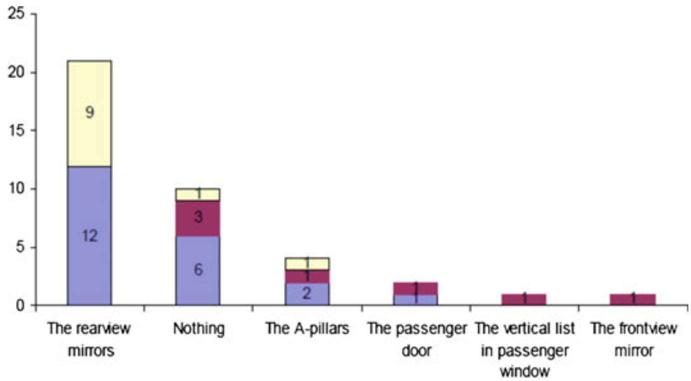


Fig. 14 Obstruction of direct vision due to body parts or attachments on a semitrailer tractor [20]

In Fig. 15 vision obstruction caused by mirrors and A-pillar while dynamic vision situations are presented from the driver's point of view [90, p. 126]. In the left picture, the left main mirror hides the traffic situation ahead within a roundabout passage. Due to the large bending angle between tractor and trailer the trailer blocks the indirect view of the past traffic situation. The adherence to the prescribed lane of the trailer cannot be controlled from this eye point position, since the position of the rear trailer axle is not shown in relation to the lane markings in the mirror. In the right picture the vision obstruction by the passenger side main and wide-angle mirror in a crossing situation is visualized. The large blind spot, caused by the high assembly of the cabin (Fig. 12) is further enhanced by the mirror-pillar combination. Thus, the behavior of other vehicles that are present in this area can't be monitored safely. Minor road users such as pedestrians, cyclists or motorcyclists can be even completely ignored due to this vision obstruction [17, p. 91], [18, p. 41].

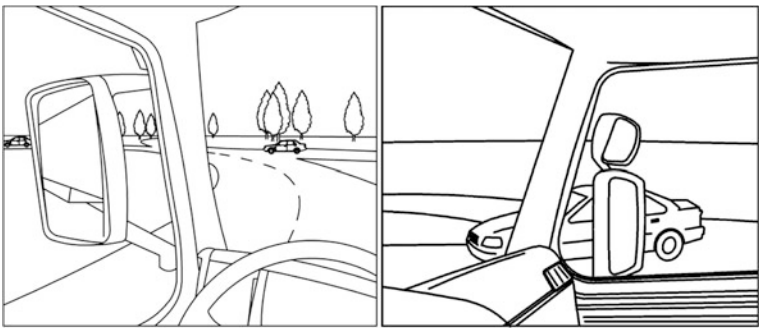


Fig. 15 Vision obstruction while round about passage (*left*) and in crossing situation (*right*) [90, p. 126; p. 132]

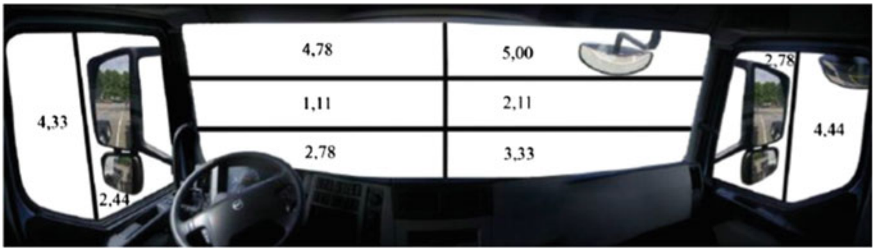


Fig. 16 Relative weighting of different areas of direct vision in trucks [20]

For a systematic evaluation of obstruction areas masking direct vision they need to be prioritized depending on the impact of their obstruction on fulfilling the driving task [64, p. 13], [76, p. 74]. Danielsson proposes relative weighting of the vision sectors for medium and heavy commercial vehicles illustrated in Fig. 16 [20, p. 59]. The weighting scale goes from 1 to 5, wherein the vision zones with high priority to fulfill the task of driving are described with low numerical values. By using a user-oriented weighted rating of the position and form of necessary components masking the vision those can be ergonomically optimized. The main and wide-angle mirrors are located in high-priority vision areas.

As shown in Figs. 12 and 15, the trailer of a semitrailer-tractor combination is in vision situations with high bend angles both target and obstruction of indirect vision (Sect. 1.1.4; [4, p. 1638], [78, p. 1]). The mirror vision is additionally restricted by the limitation of the mirror surface, the mirror curvature and the limited mounting options. Also compensation movements carried out by the driver [37, p. 21] to improve the direct and indirect vision are limited because of the human musculoskeletal system and the requirements for fulfilling the dynamic driving task.

It may be noted that the vision task in the commercial vehicle must consider particularly dynamically complex driving situations with large truck-trailer angles [6, p. 2]. The exterior mirrors are necessary for showing indirect vision but restrict the direct vision. The trailer of a tractor-trailer combination takes an exceptional role within the restricted vision. It is a bend angle between tractor and trailer dependent dynamic vision target and obstruction of indirect vision.

2.2 Design of Camera-Monitor Systems

A camera-monitor system (CMS) for the replacement of vehicle mirrors of Class I to IV [84, par. 15.2.4.2] consists according to ISO 16505 of the system components shown in Fig. 17 [90, p. 22]. Specifically, these are a device for image capturing

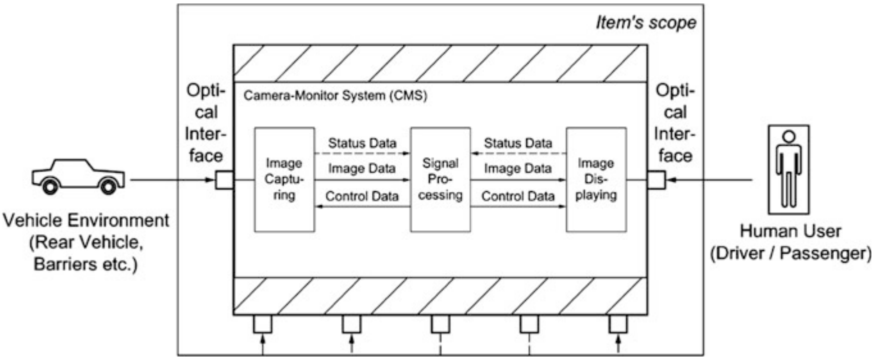


Fig. 17 System components of a camera-monitor system [90, p. 21]

(camera), a module for signal processing (controller) and a unit for image displaying (monitor). The entire system thus has at least two optical interfaces. The camera captures the relevant fields of vision of the environment and the monitor displays the processed image information to the driver. Since the vision is of great importance to fulfill the task of driving (see Sect. 2.1.3), a camera-monitor system (CMS) that provides the relevant information of indirect vision should be designed in addition to the technical parameters after ergonomic aspects.

Over the past decades, mirror systems have been adjusted due to changed and expanded regulations [83, Annex 6], [84, par. 1.5.2.4] and ergonomic developments by the system suppliers and car manufacturers to meet the requirements of changing transportation task (see Sect. 1.1.1). The current mirror systems thus are the reference for the development of new electronic vision systems, such as the CMS (see [90] Scope).

In Sect. 2.2.1 the ergonomic validation process and the tools for evaluation and documentation of the individual design steps of a vehicle are introduced. In the next section selected properties of conventional mirror systems and camera-monitor systems (CMS) are compared. As a basis for further ergonomic design the conceptual conditions of the considered CMS are determined. The position and orientation of the camera and monitor components are defined and visualized. The parameters for showing the vision areas on the necessary display portion of the monitor can be specified for standard situations.

There are remaining parameters that can't be processed with the described design method. Satisfying the claim for a monitor image in the required image quality is not sufficiently substantiated through the alignment of the monitors on the design eye point (ocular reference point). Additional information on the monitor-related localization of relevant dynamic eye point positions is required. There are also no adequate findings regarding the vision arrears of indirect vision used by the driver in dynamic vision situations.

2.2.1 Ergonomic Design Process

For ergonomic design of new vehicles and vehicle components, these have to be adapted to the abilities and skills of people [62, p. 969]. At the beginning of the ergonomics process of new vehicle concepts the customer-benefit criteria and the related ergonomics goals are to be defined within the strategic phase [3, p. 30], [52, p. 9], [73, p. 3]. In this early phase of the project, the postures and eye point positions of the defined user portfolio [3, p. 78] [61, p. 5] are needed for the related ergonomic basic design of the dimensional concept [53, p. 1] and to confirm the vehicle proportions [73, p. 4].

A data base can be collected using subject tests in parameterizable seat boxes [3, p. 19], [29, p. 49], [64, p. 19], in virtual driving simulators, as the VR environment CAVE [64, p. 16], [3, p. 23] or in real driving test [64, p. 6], [76, p. 53], [78, p. 5]. Due to the early stage of the project, surface and packaging data, as well as the production-based technology components are not available in sufficient extent. For this process step, the human model-based simulation of sitting postures has been established [38, p. 11], [61, p. 4], [66, p. 1], [67, p. 1], [73, p. 2], in addition to the proven but limited tools for the dimensional concept, as the body outline template [81] or SAE-eye ellipsoid [91]. Both for the implementation [57, p. 7] as well as for the validation [53, p. 119] of model based posture simulations representative eye point distributions are needed.

In the further course of concept development, the engineering and design-driven conditions are concretized and thereby the dimensional concept is sharpened [3, p. 30], [52, p. 10], [73, p. 6]. The ergonomic verification of the following milestones is thus based on an optimized data basis and increases, as the entire project, in extent and complexity. In addition to the simulation of realistic driving postures other issues such as the implementation and validation of intuitive interfaces and of the holistic vision evaluation are gaining in importance. Information is required about the context of use, the associated eye point positions in connection with the gaze orientation [3, p. 80], [66, p. 92] and the dynamic vision behavior of the vehicle occupants [64, p. 6]. At the latest on the ergonomics release the ergonomic vision parameters can be validated on the prototype [3, p. 30] and the simulation quality of the previous validation steps can be checked. After the series launch of the new vehicle comprehensive vision data records can be collected in real driving conditions as a basis for the initial design of the successor vehicle and the experience-based method development.

The evaluation of vision aspects [37, p. 10], [58, p. 21], [64, p. 4], [76, p. 7] plays an important role [3, p. 37] during the ergonomic verification [3, p. 30], [61, p. 20]. In conventional vehicle concepts it can be drawn between the direct vision from the vehicle, the indirect vision via mirror systems and the view on displays [58, p. 64], [76, p. 14]. For a passenger car it is differentiated within the indirect vision, between the vision via interior mirrors and the vision via exterior mirrors. In a conventional long-distance transport vehicle there is no interior mirror available, therefore only indirect vision via exterior mirrors or electronic vision systems exists.

2.2.2 Technology-Related Design Parameters

In Table 1 selected parameters for the design of a camera-monitor system (CMS) are shown. Compared to conventional mirror system, individual parameters can be varied independently of each other in a CMS. Concerning the requirement that objects appear in the CMS at least as large as in the reference mirror system ($M_{mirror} < M_{CMS}$; [90, p. 49]) the Magnification ($M_{CM}S$), the Visualization Surface on the Monitor ($A_{monitor}$) and the Displayed Vision Area ($\alpha_{monitor}$) can be designed freely. For technological reasons, there is no optical link between the design eye point (ORP) and the vision object. The Captured Vision Area (α_{camera}) can be chosen independently, taking into account all fields of vision necessary for dynamic vision situations. With a camera-monitor system (CMS), the driver-initiated head movements do not result in panning of the displayed vision area. For an ergonomic design of a CMS the required vision area should be panned automatically.

2.2.3 Defining the Conceptual Design Parameters

As a basis for the further ergonomic design of a CMS, the conceptual design parameters of the considered system are determined. For a detailed conception and evaluation of CMS approaches see the work from Fornell-Fagerström and Gardlund [27, p. 55]. The determined component positions must be evaluated in relation to the reference semitrailer-tractor combination (Fig. 18) during dynamic vision tasks (Sect. 2.1.3).

The main requirements for the positioning and orientation of the camera components are listed in Table 2. Based on these requirements, the positions of the cameras are determined as shown in Fig. 19. The cameras are mounted on the cab roof above the doors. The lenses are aligned on the respectively required field of vision [84, par. 15.2.4.2].

The main requirements for the positioning and orientation of the monitor components are listed in Table 3. Based on these requirements, the positions of the monitors are determined as shown in Fig. 20. The monitors are fixed to the A-pillars in the interior of the cab and orthogonal aligned to the design eye point (ORP) [90, p. 6]. This positioning reduces the restriction of the direct vision to a minimum.

The entire system consisting of cameras, signal processing and monitors must meet the magnification required in the ISO 16505 [90, p. 49] and show the fields of vision defined in ECE-R 46 [84 par. 1.5.2.4] (see Table 4). Thus, the fields of vision necessary to fulfill the standard situations with axial orientated tractor and

Table 1 Technology-related design parameters of a CMS

Category	Conventional Mirror System	Camera-Monitor System (CMS)
Magnification (Magnification factor between captured and displayed objects)	**M_{mirror}** • Distance between ORP and mirror surface (a_{mirror}), mirror curvature (r_{mirror}), distance between mirror surface and vision object (d_{object}), and angle of incident and emergent line of sight (β) dependent [90, p. 13]	**M_{CMS}** • Distance between ORP and display surface ($a_{monitor}$) and internal zoom factor of the CMS dependent **→M_{CMS} can be chosen with the preset M_{CMS} >= M_{mirror}** [90, p. 49]
Visualization Surface	**A_{mirror}** • Magnification (M_{mirror}) and displayed field of vision (at least ECE-R46) dependent [84, par. 1.5.2.4] [90, p. 13]	**$A_{monitor}$** • Magnification (M_{CMS}) and displayed field of vision (at least ECE-R46) dependent **→$A_{monitor}$ can be chosen with the preset M_{CMS} >= M_{mirror}** [84, par. 1.5.2.4] [90, p. 49]
Captured Vision Area	**α_{mirror} = Displayed Vision Area** • Magnification (M_{mirror}) Visualization Surface (A_{mirror}) dependent [90, p. 13]	**Captured Vision Area = $\alpha_{camera} \neq \alpha_{monitor}$** **→Captured Vision Area can be chosen, taking into account all necessary fields of vision**
Displayed Vision Area	**α_{mirror} = Captured Vision Area** • Magnification (M_{mirror}) Visualization Surface (A_{mirror}) dependent [90, p. 13]	**Displayed Vision Area = $\alpha_{monitor} \neq \alpha_{camera}$** **→Displayed Vision Area can be chosen with the preset M_{CMS} >= M_{mirror}** [90, p. 49]
Optical Link	• Optical Link between eye point, mirror surface and vision object • Displayed Vision Area = Captured Vision Area • Vision area can be panned by driver-initiated head movement	• No optical link between eye point and vision object • Displayed Vision Area ≠ Captured Vision Area • Vision area can't be panned by driver-initiated head movement **→Required vision area should be shown automatically**

Fig. 18 Reference semitrailer-tractor combination to evaluate the mounting positions of the camera-monitor system (CMS) [19]

Table 2 Requirements for camera position and orientation of a CMS

CMS Component	Requirement
Camera	• Maximum CO_2 reduction is to be achieved [21, p. 59] • Minimal wind noise is to be achieved [33, p. 536] • Minimal soiling of the camera and the vehicle is to be achieved [44, p. 720] • Capture of the required fields of vision (min. [84, par. 1.5.2.4] →**Camera position and orientation can be determined** [27, p. 71]

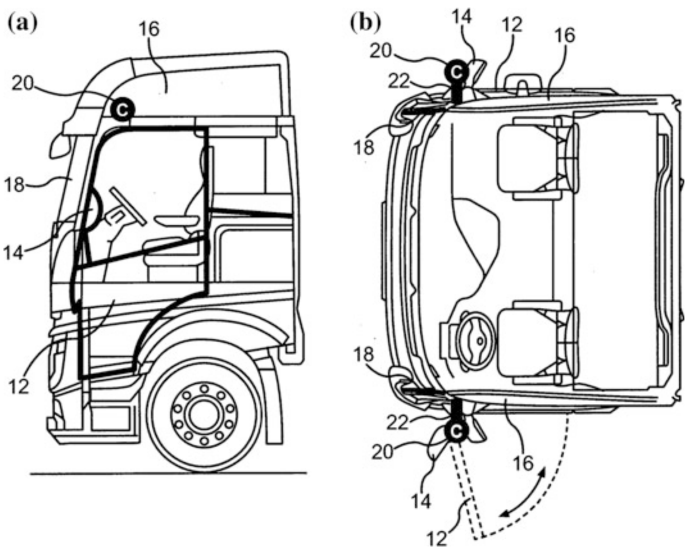


Fig. 19 Determined camera position [10, p. 5]

Table 3 Requirements for monitor position and orientation of a CMS

CMS Component	Requirement
Monitor	• Monitors have to be meaningful right arranged for the driver (left portion for FOV driver's side, right portion for FOV passenger side) [90, p. 95] • Direct view onto the monitor must not be obstructed by other components of the vehicle [90, p. 51] • Restriction of direct vision from the vehicle should be reduced to a minimum [90, p. 51] • Display surface should be orientated to comply with the required optimal image quality for the design eye point (ORP) [90, p. 68] • **Monitor position and orientation can be determined** [27, p. 75]

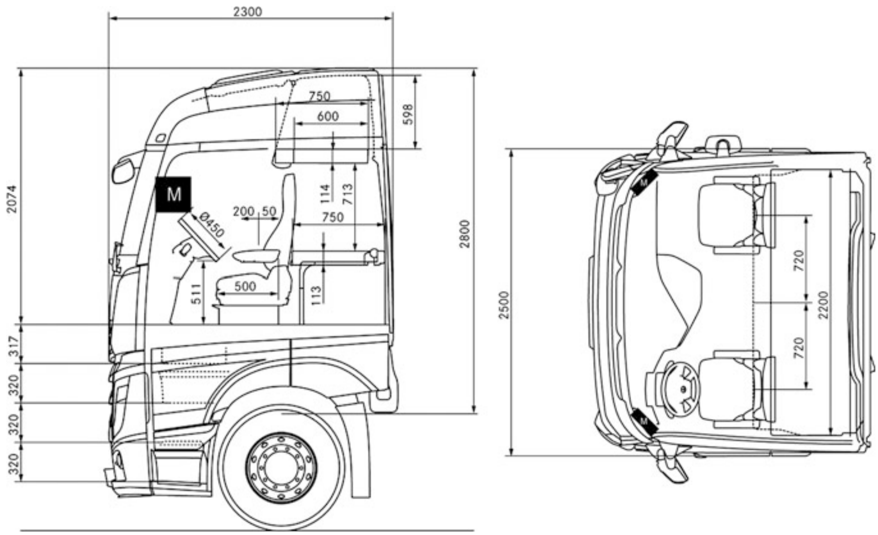


Fig. 20 Determined monitor position [10, p. 5]

Table 4 Requirements for the entire camera-monitor system (CMS)

CMS Component	**Requirement**
Entire System	• Conformance to the required magnification [90, p. 49] • Presentation of the required fields of vision (min. [84, par. 1.5.2.4] →**Displayed vision area (standards situation) on necessary display surface can be determined**
Open issues	• **For all drivers a CMS image with the required image quality must be represented in all relevant driving situations.** [90, p. 68] • **The fields of vision, needed by the driver in “special driving situations” have to be shown automatically.** [90, p. 93]

semitrailer represented on the display area can be specified. The fulfillment of the claim for a monitor display with the required image quality for all drivers in all driving situations is not sufficiently substantiated through the alignment of the monitors on the design eye point (ORP). Additional information on the monitor-related localization of relevant dynamic eye point positions is required. There are also no adequate findings regarding the vision arrears of indirect vision used by the driver in dynamic vision situations.

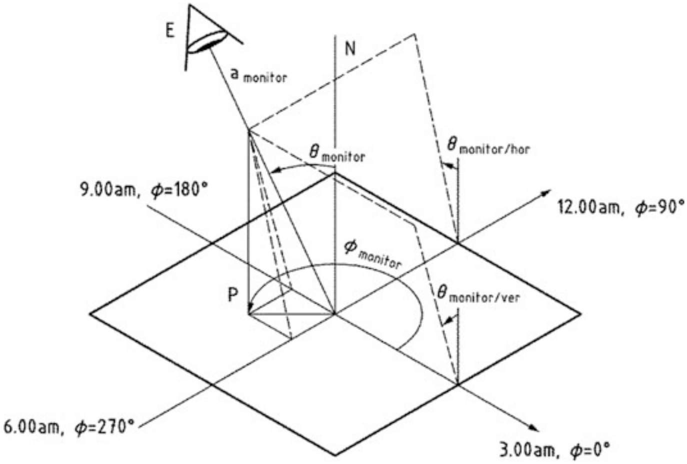


Fig. 21 Polar coordinate system of the monitor [90, p. 22]

2.2.4 Viewing Angle-Dependent Display Quality of the Monitor

Concerning the viewing angles ($\Phi_{monitor}$) and ($\theta_{monitor}$) of the monitor polar coordinate system (see Fig. 21) and the distance between monitor and eye point ($a_{monitor}$) each eye point position of the driver relative to the monitor installed in the vehicle can be discribed [90, p. 22].

The display quality of the monitor depends on the viewing angles ($\Phi_{monitor}$) and ($\theta_{monitor}$) (see Fig. 22; [90, p. 68]). Therefore, all other relevant eye point positions next to the design eye point ORP [90, p. 6] must be taken into account for the ergonomic alignment of the monitor. In particular, dynamic eye point positions, that describe the spread of eye point distributions within the "special driving situations" [90, p. 93] are of interest. These dynamic eye point positions are depending on the body dimensions of the drivers, the vehicle variant and the vision dependent [40, p. 5], [70, p. 4], [90, p. 107].

2.2.5 Situation-Dependent Adaptation of the Displayed Vision Area

Within the informative Annex B of ISO 16505 [90, p. 97] an extension of the displayed vision area to fulfill the driving task in "special driving situations" such as turning maneuvers, roundabouts, crossings and maneuvering [90, p. 125] is proposed [90, p. 109]. In Fig. 23, a proposal of extended displayed vision areas is shown. This proposal is based on an evaluation of the result of values [90, p. 109] collected by Blomdahl [2, p. 26] regarding average and maximum head movement during "special driving situations".

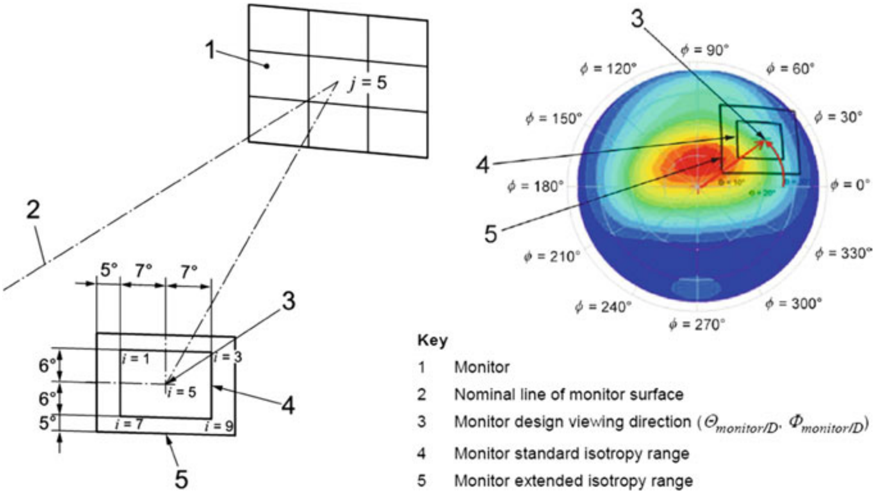


Fig. 22 Measurement of the viewing angle-dependent display quality of the monitor [90, p. 69]

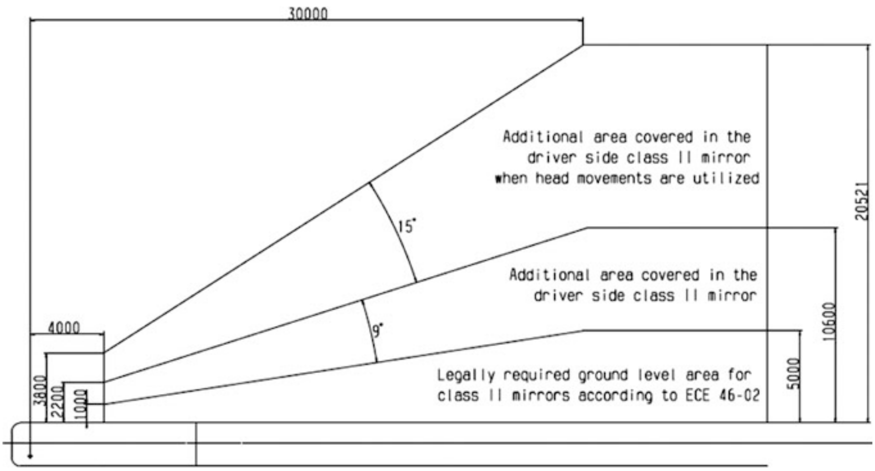


Fig. 23 Proposal for the expansion of the displayed vision area for "special driving situations" [27, Appendix D]

As shown by Fornell-Fagerström and Gardlund [27, p. 49], the presentation of this extended vision areas with the magnification factor required in ISO 16505 leads to very large displays that hinder the direct line of sight in a non-tolerable extent. On closer examination on the vision interaction of drivers with conventional mirrors (Sect. 1.1.4) it is noticeable that the drivers does not expand the field of vision by the head movement but pan it temporarily [4, p. 1638], [9, p. 298], [27, p. 53],

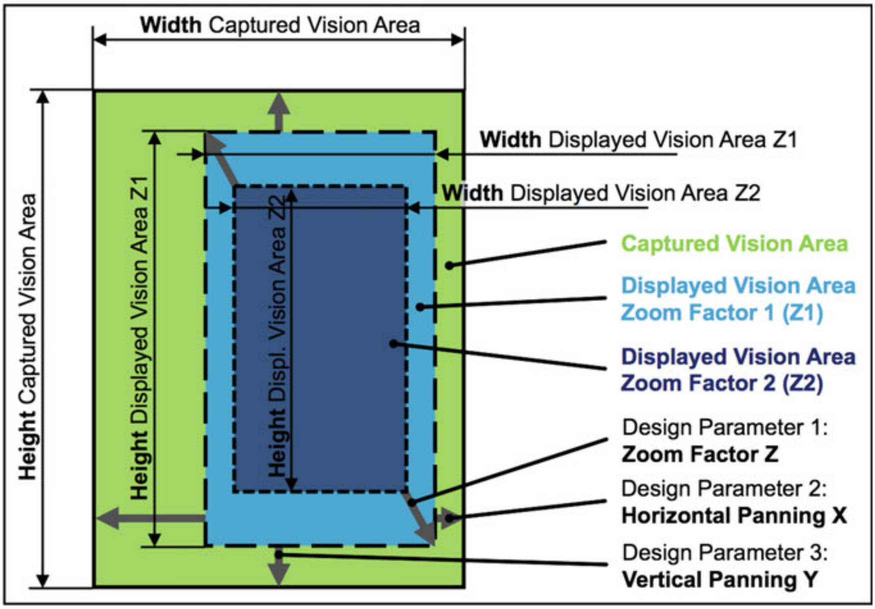


Fig. 24 Derivation of the design parameters of the displayed vision areas of a CMS

[90, p. 111]. The design parameters necessary for dynamic panning of to the displayed visual area [8, p. 1] are summarized in Fig. 24.

The camera captures a sufficient vision area for providing image information in all driving situations. A section of the captured vision area will appear as the displayed vision area on the monitor. Width and height of the displayed vision area can be adjusted within the boundary conditions of Table 1 by changing the zoom level (Z). The position of the displayed vision area is varied by the horizontal (X) and vertical (Y) panning [8, p. 1].

The automated adjustment of the displayed vision area can be implemented by a vehicle parameter-dependent control and stored in a panning characteristic. There are parameters, such as the vehicle speed, the steering wheel angle and the calculated bend angle proposed for defining the necessary displayed vision area [8, p. 9].

In Fig. 25 situation dependent simulations of displayed vision areas of a CMS are shown. The switch from vision situation 1 to vision situation 2 occurs via a panning of the selected portion within the captured vision area. In order to implement within a real system, a description of the vision situation dependent and automatically controllable by vehicle parameters, a driver required dynamic field of vision is needed. Thus, the definition of a characteristic curve for situation-dependent representation of displayed dynamic vision areas within a CMS is possible.

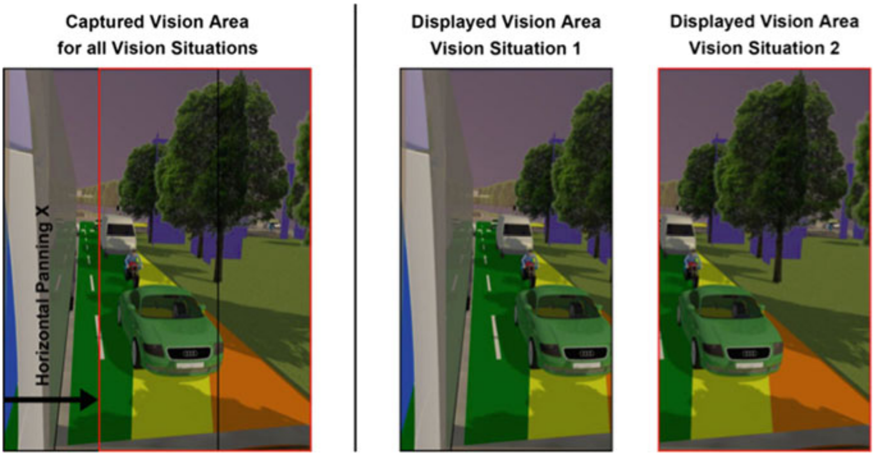


Fig. 25 Simulation of the situation-dependent panning of the displayed vision area [27, p. 53]

3 Ergonomic Design of Camera-Monitor Systems

The findings obtained from the analysis of the dynamic vision situations [11, p. 85] are used for the ergonomic design of remaining issues in the design of camera-monitor systems (CMS). The viewing angle-dependent display quality of monitors is considered by aligning on the dynamic eye point positions. The automated adjustment of the displayed vision areas of indirect vision is implemented through a panning characteristic of the CMS. Thus, the advanced vision system is adapted to the ergonomic needs of truck drivers.

3.1 Monitor Alignment on the Dynamic Eye Point Positions

After analyzing the static and dynamic eye point distributions it was found that the two comparison groups differ relevantly in the position of the mean values and in the spread of the distribution [11, p. 87]. Therefore, dynamic eye point distributions have to be used for the ergonomic design of vehicle components that are integrated in a visual interaction.

All dynamic eye point positions, collected during long-distance transport use cases, are summarized in a three-dimensional body. This design body is called eye-ellipsoid and includes at least 90 % of the dynamic eye point positions when looking into the evaluated areas of interest (AOI). Thus the eye ellipsoid represents the design-relevant portion [70, p. 1] of the dynamic eye point positions. The eye-ellipsoid generated for the reference vehicle is shown in Fig. 26. In addition to the design body the center point (mean dyn eye point total) and the middle eye point positions of n = 46 subjects (mean dyn eye point subject) are displayed.

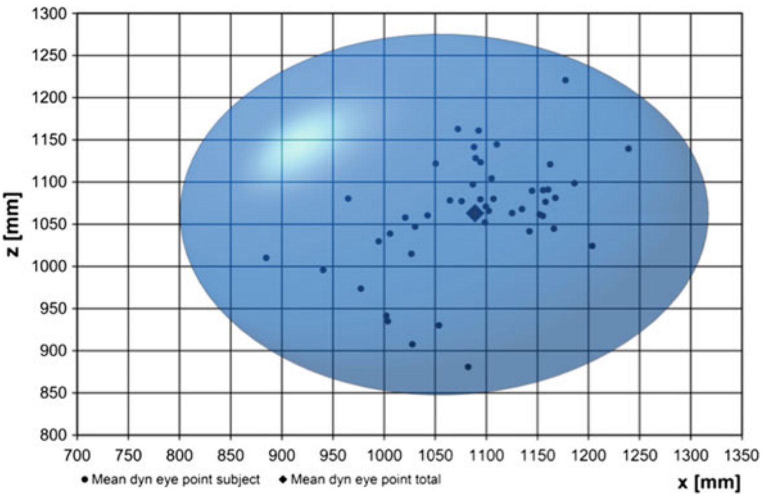


Fig. 26 Eye-ellipsoid to describe the dyn eye point positions n = 46 side view [11]

The image quality of a display used in the CMS depends on the angle between gaze direction and display surface (Sect. 2.2.4). Display-dependent opening angle are defined which ensure a sufficient image quality. These display-opening angles are dependent on the technology and design of the monitor, and are given as horizontal and vertical angle values.

In Fig. 27 the driver side monitor of the CMS and the eye-ellipsoid is shown inside the cabin. For ergonomic design, the display surface is aligned to the mean of the eye-ellipsoid. The attachment to the cab is chosen so that the restriction of direct vision caused by the system (Sect. 2.1.3) is minimal. Since the center of the design body is the mean of all considered dynamic eye point distributions it is ensured that the monitor is optimally positioned for the dynamic vision interaction in the vehicle. For checking the selected monitor position and orientation it must be considered that the entire eye-ellipsoid is enclosed by the display-opening angles. It is thus ensured that for at least 90 % of the dynamic eye point positions an optimal image quality is shown on the monitor. The design steps for positioning the driver side monitor are also performed for the passenger side. The monitor the CMS illustrated in Fig. 27 is thus designed ergonomically to the real usage of the driver in the vehicle.

3.2 *Panning Characteristic of the Camera-Monitor Systems*

In a camera-monitor system (CMS) as described in Sect. 2.2, the displayed vision areas of indirect vision can be adjusted automatically according to the driver's situational needs [8, p. 9]. For the horizontal field of vision, the spread of the

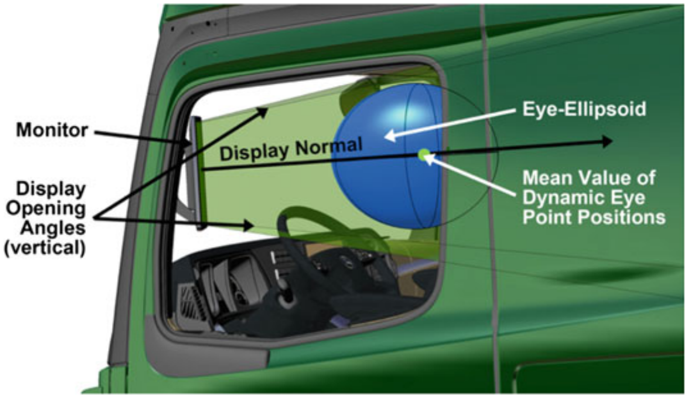


Fig. 27 Alignment of the display surface to the mean value of the dynamic eye point positions and checking of the opening angle based on the eye-ellipsoid [11]

displayed image is defined by the inner and outer indirect vision angle (Fig. 28). By defining this angle on a truck-trailer angle range of −90° to +90° the horizontal panning characteristics is created. The area between the inner and outer curve represents the displayed dynamic vision area. The possibility of variation of the displayed vision area is limited by the opening angle of the camera. For reference, the regulatory requirements of ECE-R46 [84, par. 15.2.4.2] are plotted for the static truck-trailer combination. The distance between the inner and outer indirect vision angles affects the object size and the width of the displayed vision area. If the distance increases larger areas of the vehicle environment can be displayed, the display size of objects such as other vehicles or track marks is thereby reduced. The distance may be chosen freely when considering the legal requirements [90, p. 49].

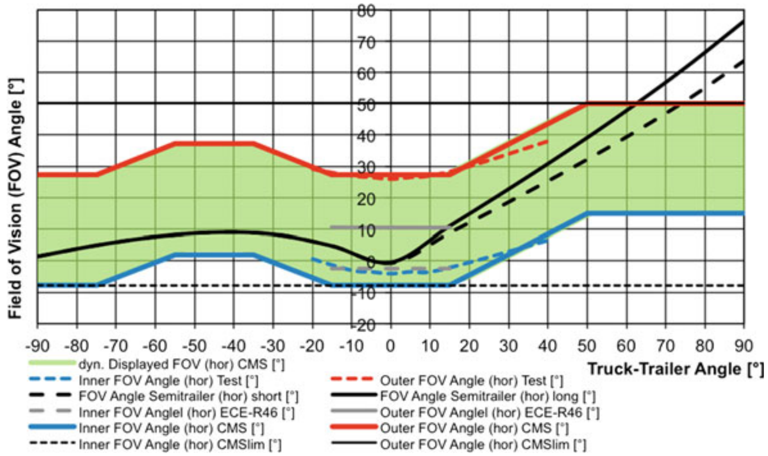


Fig. 28 Horizontal panning characteristics of CMS class II driver side [11]

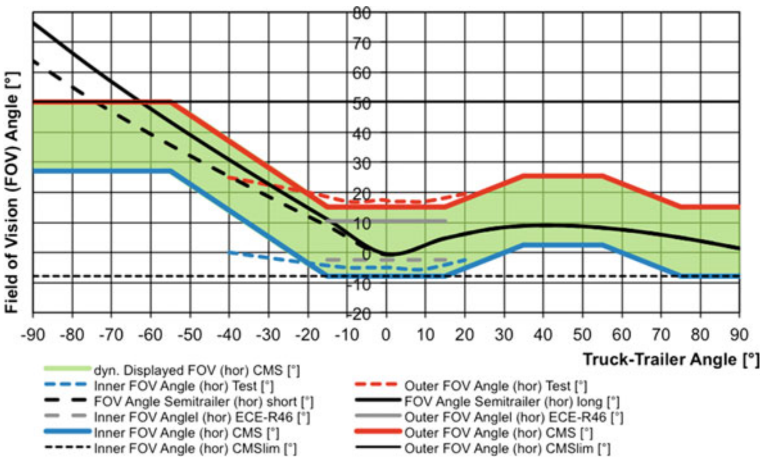


Fig. 29 Horizontal panning characteristics of CMS class II passenger side [11]

The findings from the analysis of the dynamic vision areas enable the ergonomic design of the panning characteristics for camera-monitor systems (CMS) [11, p. 11]. In Fig. 28 the inner and outer FOV angle collected during the experiment are drawn in the truck-trailer angle range between −20° and +40°. This is the bend angle range used in the left main mirror during the evaluated visual situations. It represents the typical long-distance use of a commercial vehicle. The panning characteristics are created with a constant course between −15° and +15° according to the FOV angles of the experiment. For the remaining truck-trailer angles the displayed vision area is panned. In Fig. 29 the horizontal panning characteristics for the passenger side is shown.

In order to interpret the characteristics for the non-appeared bend angle ranges, which are only used during maneuvering, the vision relevant semitrailers behavior is considered. Within maneuvering especially the rear and front edges of the semitrailer are of interest [9, p. 298], [78, p. 3]. Therefore, the indirect vision area angle dependent on positions of the relevant semitrailer edges for short and long trailers are shown in Fig. 28. When turning left (bend angle >10°), the rear semitrailer edge is required by the driver. When turning right (bend angle <−10°) the front edge semitrailer is relevant. Under the requirement that the relevant semitrailer edges are displayed as long as possible by the camera-monitor system (CMS), the panning characteristics is defined across the entire bend angle range.

The procedure described for the ergonomic design of the panning characteristics is also performed for the passenger-side displayed vision areas. The findings gained through the analysis of the dynamic interaction between driver's gazes, conventional mirror system and vision targets enables the ergonomic design of technology-related necessary dynamic display behavior of a CMS. Thus, a camera-monitor system (CMS), which takes the needs of drivers in addition to the legal requirements into consideration, can be developed.

4 Conclusion

Caused by the potential for fuel savings and the technology-related benefits for the truck drivers, it is of interest to replace conventional mirror systems of a commercial vehicle with camera-monitor systems (CMS). The electronic vision system is directly involved in the dynamic viewing interaction between driver, vehicle and environment and therefore needs to be designed according to ergonomic methods. For this purpose, information about the dynamic vision behavior of the truck driver while using conventional mirrors is needed. Firstly, there is the need for information about the driver's dynamic indirect vision areas to fulfill typical long-distance transport tasks. Secondly, the dynamic eye point scatter plot defined by the driver usage during dynamic vision situations.

The truck-trailer-angle is identified as a predictor of the dynamic vision behavior in heavy-duty vehicles. To determine the dynamic indirect fields of view (FOVs) the vision data is therefore evaluated with a calculation model based on the truck-trailer-angle. This results in vision angle curves representing the used dynamic indirect fields of view of truck drivers in the evaluated vision situations. A dependence of these curves from the dynamic vision situations is shown.

The findings of the analysis of dynamic vision situations are used for the ergonomic design of a camera-monitor system (CMS). The viewing angle-dependent image quality of the monitors is taken into consideration for the display adjustment depending on the dynamic eye point positions. The automated adaptation of the shown dynamic fields of vision is implemented through a truck-trailer-angle-dependent panning characteristic of the CMS. Thus, the advanced vision system is adapted to the ergonomic needs of truck drivers. The knowledge gained can be used for the ergonomic design of further vision-relevant systems. The optimization of driving posture and movement simulation is a further application of the recorded data and the obtained findings.

References

1. Abendroth, B., Bruder, R.: Die Leistungsfähigkeit des Menschen für die Fahrzeugführung. In: Winner, H., et al. (Hrsg.) Handbuch Fahrerassistenzsysteme, Vieweg +Teubner Verlag, Springer Fachmedien, Wiesbaden (2012)
2. Blomdahl, P.: Mirror usage in trucks. In: Fornell-Fagerström, K., Gardlund, A. (eds.) Mirror replacement in trucks. Master Thesis, Chalmers University of Technology, Gothenburg (2012)
3. Bothe, A.: Entwicklung eines parameterbasierten CAD-Modells zur Ergonomieabsicherung einer Mittelklasselimousine. Universität Stuttgart, Stuttgart, Thesis (2010)
4. Bothe, A., Wohlfarth, E., Bruder, R.: Dimensioning of actual fields of indirect vision of commercial vehicles by analyzing dynamic vision situations. In: Conference Proceedings 21. Aachener Kolloquium Fahrzeug- und Motorentechnik, Aachen/Beijing (2012a)
5. Bothe, A., Wohlfarth, E., Bruder, R.: Analysis of dynamic vision situations as a basic condition for human model based design of commercial vehicles. In: Human Solutions RAMSIS User Conference 2012, Hamburg (2012b)

6. Bothe, A., Wohlfarth, E., Bruder, R.: Input from the method of analyzing dynamic vision situations into the ISO working group regarding “Ergonomic and performance aspects of Camera-Monitor Systems”. ISO Working Group Conference, Yokohama (2012c)
7. Bothe, A., Wohlfarth, E., Kirschbaum, M., Bruder, R.: Selection of relevant test parameters and their influence on first results of the analysis of dynamic vision situations in commercial vehicles, 59. Kongress der Gesellschaft für Arbeitswissenschaft (GfA), Krefeld (2013a)
8. Bothe, A., Wohlfarth, E., Kirschbaum, M., Bruder, R.: Variation of actual fields of indirect vision in commercial vehicles during dynamic vision situations. In: 13th Stuttgart International Symposium of Automotive and Engine Technology, Stuttgart (2013b)
9. Bothe, A., Kober, W., Wohlfarth, E.: Verfahren zum Steuern eines Kamerasystems für ein Fahrzeug und Kamerasystem. Patent Application, Deutsches Patent- und Markenamt, München (2013c)
10. Bothe, A., Wohlfarth, E.: Nutzfahrzeug mit Bilderfassungs einrichtung. Patent Application, Deutsches Patent- und Markenamt, München (2014)
11. Bothe, A.: Analyse dynamischer Sichtsituationen zur ergonomischen Auslegung von Kamera-Monitor-Systemen (KMS) in schweren Nutzfahrzeugen, Dissertation, Darmstadt (2015)
12. Brähler, H.: Konzeption von Nutzfahrzeugen. In: Hoepke, E., Breuer, S. (eds.) Nutzfahrzeugtechnik. Springer Fachmedien, Wiesbaden (2010)
13. Braun, H., Kolb, G.: LKW Ein Lehrbuch und Nachschlagewerk. Kirschbaum Verlag, Bonn (2008)
14. Breuer, S., Kopp, S.: Fahrmechanik. In: Hoepke, E., Breuer, S. (eds.) Nutzfahrzeugtechnik. Springer Fachmedien, Wiesbaden (2013)
15. Bubb, H.: Analyse der Geschwindigkeitswahrnehmung im Kraftfahrzeug. Zeitschrift für Arbeitswissenschaft, vol. 31. Ergonomia Verlag, Stuttgart (1977)
16. Cohen, A.: Visuelle Orientierung im Straßenverkehr. Schweizerische Beratungsstelle für Unfallverhütung, Bern (1998)
17. Cook, S., et al.: The Development of Improvements to Drivers’ Direct and Indirect Vision from Vehicles—Phase 1. Loughborough Design School, Loughborough (2010)
18. Cook, S., et al.: The Development of Improvements to Drivers’ Direct and Indirect Vision from Vehicles– Phase 2. Loughborough Design School, Loughborough (2011)
19. Daimler AG (2014): Global Media Site. http://media.daimler.com/dcmedia/home/de
20. Danielsson, H.: Critical vision and Adapted driving postures. In: Fornell-Fagerström, K., Gardlund, A., (2012): Mirror replacement in trucks. Master Thesis, Chalmers University of Technology, Gothenburg (2011)
21. Devesa, A.; Indinger, T.: Verbrauchsreduktion an Nutzfahrzeugkombinationen durch aerodynamische Maßnahmen. In: FAT-Schriftenreihe Nr. 237. Forschungsvereinigung Automobiltechnik e.V., Berlin (2011)
22. Statistisches Bundesamt. Daten zur Energiepreisentwicklung- Lange Reihen von Januar 2000 bis Februar 2014. Statistisches Bundesamt, Wiesbaden (2014)
23. Donges, E.: Aspekte der Aktiven Sicherheit bei der Führung von Personenkraftwagen. Automobil-Industrie, vol. 27, Vogel Business Media, Würzburg (1982)
24. Donges, E.: Fahrerverhaltensmodelle. In: Winner, H., et al. (Hrsg.): Handbuch Fahrerassistenzsysteme, Springer Fachmedien, Wiesbaden (2012)
25. Engelkamp, J., Zimmer, H.: Lehrbuch der Kognitiven Psychologie. Hogrefe, Göttingen (2006)
26. Ernst, C.-S., Olschewski, I., Eckstein, L.: Implikation der neuen CO_2-Gesetzgebung für schwere Nutzfahrzeuge im Fernverkehr. In: Göpfert, I., et. al. (Hrsg.), Automobillogistik—Stand und Zukunftstrends, Springer Gabler, Wiesbaden (2013)
27. Fornell-Fagerström, K., Gardlund, A.: Mirror replacement in trucks. Master Thesis, Chalmers University of Technology, Gothenburg (2012)
28. Frank, T.: Aerodynamik von schweren Nutzfahrzeugen—Stand des Wissens. In: FAT-Schriftenreihe Nr. 241, Forschungsvereinigung Automobiltechnik e.V., Berlin (2012)
29. Fröhmel, C.: Validierung des RAMSIS-Krafthaltungsmodells. Dissertation. Technische Universität München, Lehrstuhl für Ergonomie, München (2010)

30. Göpfert, I., Schulz, M., Wellbrock, W.: Trends in der Automobillogistik. In: Göpfert, Ingrid, Braun, David, Schulz, Matthias (eds.) Automobillogistik—Stand und Zukunftstrends. Springer Fachmedien, Wiesbaden (2013)
31. Grandjean, E.: Physiologische Arbeitsgestaltung– Leitfaden der Ergonomie. Ott-Verlag, Thun (1991)
32. Hakuli, S., et al.: Kooperative Automation. In: Winner, H., et al. (Hrsg.) Handbuch Fahrerassistenzsysteme, Springer Fachmedien, Wiesbaden (2012)
33. Helfer, M.: Umströmungsgeräusche. In: Schütz, T. (ed.) Hucho—Aerodynamik des Automobils. Springer Fachmedien, Wiesbaden (2013)
34. Herrmann, H.-J., Scheuer, S.: Optische Anforderungen und Prüfverfahren für Kamera-Monitor-Systeme für indirekte Sicht in Fahrzeugen. INS Zwischenbericht, TÜV Rheinland LGA Products GmbH, Köln (2011)
35. Hoepke, E., Brähler, H.: Einführung in die Nutzfahrzeugtechnik. In: Hoepke, E., Breuer, S. (eds.) Nutzfahrzeugtechnik. Springer Fachmedien, Wiesbaden (2010)
36. Hoffman, J.: Visual attention and eye movements. In: Pashler, H. (ed.) attention. University College London Press, London (1998)
37. Hudelmaier, J.: Sichtanalyse im Pkw unter Berücksichtigung von Bewegung und individuellen Körpercharakteristika. Dissertation. Technische Universität München, Lehrstuhl für Ergonomie, München (2002)
38. Human Solutions. RAMSIS in CATIA V5– Handbuch Version 3.8.37. Human Solutions GmbH, Kaiserslautern (2013)
39. James, W.: The principles of psychology. American Science Series—Advanced Course, vol. I. Harvard University, Henry Holt and Company, New York (1890)
40. Jürgens, H., Helbig, K., Küchmeister, G.: Anthropometrische und biomechanische Untersuchungen an Fahrern und Fahrerinnen von Reisebussen und Lastkraftwagen– Im Auftrag der Berufsgenossenschaft für Fahrzeughaltungen. Forschungsgruppe Industrieanthropologie der Universität Kiel, Kiel, Projektbericht (2000)
41. Karsh, R., Breitenbach, F.W.: Looking at looking: The amorphous fixation measure. In: Groner, R., et al. (eds.) Eye Movements and Psychological Functions: International Views. Erlbaum, Hillsdale (1983)
42. Kasties, G.: Telematik. In: Braess, H.-H., Seiffert, U. (eds.) Vieweg Handbuch Kraftfahrzeugtechnik. Springer Fachmedien, Wiesbaden (2013)
43. Koch, S.: Logistik—Eine Einführung in Ökonomie und Nachhaltigkeit. Springer, Berlin (2012)
44. Kopp, S., Frank, T.: Nutzfahrzeuge. In: Schütz, T. (ed.) Hucho—Aerodynamik des Automobils. Springer Fachmedien, Wiesbaden (2013)
45. Krueger, H.: Sehen. In: Schmidtke, H., (ed.) Handbuch der Ergonomie mit ergonomischen Konstruktionsrichtlinien und Methoden. Bundesamt für Wehrtechnik und Beschaffung, Koblenz (1989)
46. Lang, W., Kunz, M.: Sichtsystem für ein Nutzfahrzeug zur Darstellung von gesetzlich vorgeschriebenen Sichtfeldern eines Hauptspiegels und eines Weitwinkelspiegels. Patent Application, Deutsches Patent—und Markenamt, München (2012)
47. Lang, W., et al.: Spiegelersatzsystem für ein Fahrzeug. Patentschrift, Deutsches Patent—und Markenamt, München (2012)
48. Lang, W., et al.: Kameraarm für eine Spiegelersatzsystem-Kamera eines Kraftfahrzeugs. Patentschrift, Deutsches Patent—und Markenamt, München (2012)
49. Luczak, H.: Untersuchungen informatorischer Belastung und Beanspruchung des Menschen. Dissertation. In: Fortschrittsberichte der VDI-Zeitschriften, Reihe 10, Nr. 2, VDI-Verlag, Düsseldorf (1975)
50. Maierhofer, B.: Gütertransport der Zukunft: Wie verändert sich das Nutzfahrzeug? In: Automobiltechnische Zeitschrift Sonderausgabe 125 Jahre Automobil, Springer Fachmedien, Wiesbaden (2011)
51. MEKRAtronics (2013): Kamera-Monitorsysteme. Product Catalog 2013, MEKRAtronics, Ergersheim

52. Meurle, J., Rigel, S., Muigg, A.: Fahrzeugeigenschaften—Erwartungshaltung des Kunden und technische Umsetzbarkeit. In: Vortragsreihe Fahrzeugtechnik, Institut für Fahrzeugtechnik, FH JOANNEUM, Graz (2009)
53. Müller, A.: Systematische und nutzerzentrierte Generierung des Pkw-Maßkonzepts als Grundlage des Interior—und Exteriordesign. Dissertation. Universität Stuttgart, Stuttgart (2010)
54. Pannasch, S.: Ereignisbezogene Veränderungen der visuellen Fixationsdauer. Dissertation. Technische Universität Dresden, Dresden (2003)
55. Partmann, T., Reinig, H.-J., Struck, G.: Blickbewegungsmessung als Werkzeug für die Gestaltung und Bewertung von bord- und straßenseitigen Informationssystemen für den Kraftfahrer. Forschungsvereinigung Automobiltechnik e.V. (FAT), Frankfurt am Main (1996)
56. Rasmussen, J.: Skills, Rules, and Knowledges; Signals, Signs, and Symbols, and Other Distinctions in Human Performance Models. In: IEEE Transactions on Systems, Man and Cybernetics, New York (1983)
57. Reed, M., Manary, M., Flannagan, C., Schneider, L.: Comparison of Methods for Predicting Automobile Driver Posture. Digital Human Modeling for Design and Engineering, Technical Paper 2000–01-2180, SAE International, Dearborn, Michigan (2000)
58. Remlinger, W.M.: Analyse von Sichteinschränkungen im Fahrzeug. Dissertation. Technische Universität München, Lehrstuhl für Ergonomie, München (2013)
59. Renschler, A.: Vision 20–20, European Automobile Manufacturers Association (ACEA) Pressekonferenz auf der IAA 2008, Hannover (2008)
60. Rockwell, T., et al.: Drivers' Eye Movements: An Apparatus and Calibration. Highway Research Record 247, National Research Council, Washington (1968)
61. Schaller, F.: Einsatz von RAMSIS in der Innenraumgestaltung bei BMW. In: Conference Proceedings RAMSIS User Conference 2008, Human Solutions, Kaiserslautern (2008)
62. Schlick, C., Bruder, R., Luczak, H.: Arbeitswissenschaft. 3. vollständig überarbeitete und erweiterte Auflage, Springer, Berlin, (2010)
63. Schober, H.: Das Sehen. Fachbuchverlag, Leipzig, II. Band (1954)
64. Scholly, U.: Sicht im PKW—Evaluationsmethoden in der Praxis. In: Conference Proceedings Herbstkonferenz der Gesellschaft für Arbeitswissenschaft, Gesellschaft für Arbeitswissenschaft (GfA), Stuttgart (2006)
65. Schweigert, M.: Fahrerblickverhalten und Nebenaufgaben. Dissertation. Technische Universität München, Lehrstuhl für Ergonomie, München (2002)
66. Seidl, A.: Das Menschmodell RAMSIS Analyse, Synthese und Simulation dreidimensionaler Körperhaltungen des Menschen. Dissertation. Technische Universität München, Lehrstuhl für Ergonomie, München (1994)
67. Seidl, A.: Computer-Menschmodelle in der Ergonomie. In: Schmidtke, H., (ed.) Handbuch der Ergonomie mit ergonomischen Konstruktionsrichtlinien und Methoden. Bundesamt für Wehrtechnik und Beschaffung, Koblenz (1997)
68. Shell. Shell Lkw-Studie. Fakten, Trends und Perspektiven im Straßengüterverkehr bis 2030. Shell Deutschland, Hamburg (2010)
69. Southall, D., Tait, R., Walsh, T.: Drivers' field of view from large vehicles. Phase 2. The Department of Environment, Transport and the Regions (DETR), London (1998)
70. Speyer, H.: RAMSIS— On the Definition and Generation of Optimal Test Samples for Design Problems. Human Solutions GmbH, Kaiserslautern (2005)
71. Tait, R., Southall, D.: Drivers' field of view from large vehicles. Phase 3. The Department of Environment, Transport and the Regions (DETR), London (1998)
72. Thom, R.: Sind Regularien die wahre Triebfeder der Entwicklung? Teil 2: Verbrauchsvorschriften und Energieressourcen. Lecture Notes. Universität Stuttgart, Stuttgart (2006)
73. Wagner, P.-O.: Ergonomie im Produktentstehungsprozess. In: Conference Proceedings RAMSIS User Conference 2007, Human Solutions, Kaiserslautern (2007)
74. Wickens, C., Hollands, J.: Engineering Psychology and Human Performance, 3rd edn. Prentice Hall, Upper Saddle River (1999)

75. Woyna, L., Gloger S., Bruder R.: Methodische Ansätze zur Sichtbeurteilung im Fahrzeug. In: Conference Proceedings 56. Kongress der Gesellschaft für Arbeitswissenschaft (GfA), Dortmund (2010)
76. Woyna, L.: Fahrerzentrierter Kriterienkatalog zur Außensichtbeurteilung im PKW. Dissertation. Technische Universität Darmstadt, Darmstadt (2013)
77. Yarbus, A.: Eye Movements and Vision. Plenum Press, New York (1967)
78. Zaindl, A., et al.: Method to analyze the body movement for the mirror view and test method to verify the simulated results. In: 2nd International Digital Human Modeling Symposium, Ann Arbor (2013)
79. Ziegler, M.: Nutzfahrzeuge für den globalen Einsatz. In: Volker Schindler; Immo Sievers (Hrsg.): Forschung für das Auto von morgen. Aus Tradition entsteht Zukunft. Springer, Berlin (2008)

Standards, Laws and Regulations

80. AISÖ Arbeitsgemeinschaft Internationaler Strassenverkehrsunternehmer Österreichs: ECO-DRIVING SAFELY für Lkw—Ökonomisch und ökologisch denken!. AISÖ, Wien (2010)
81. DIN 33408-1 Deutsches Institut für Normung e.V.: Deutsche Norm, Körperumrissschablonen —Teil 1: Für Sitzarbeitsplätze, Beuth Verlag, Berlin (2008)
82. StVZO: §35b Abs. 2: Richtlinie für die Sicht aus Kraftfahrzeugen. Verordnung straßenverkehrsrechtlicher Vorschriften, Bonn (1987)
83. ECE-R 46/01 United Nations Economic Commission for Europe: Einheitliche Vorschrift für die Genehmigung von Einrichtungen für indirekte Sicht und von Kraftfahrzeugen hinsichtlich der Anbringung solcher Einrichtungen. UNECE, Geneva (2005)
84. ECE-R 46/02 United Nations Economic Commission for Europe: Einheitliche Vorschrift für die Genehmigung von Einrichtungen für indirekte Sicht und von Kraftfahrzeugen hinsichtlich der Anbringung solcher Einrichtungen. UNECE, Geneva (2009)
85. EGV Nr. 561 Europäisches Parlament und Rat: Verordnung (EG) Nr. 561/2006 zur Harmonisierung bestimmter Sozialvorschriften im Straßenverkehr…, In: Amtsblatt der Europäischen Union L102/1, Amt für Veröffentlichungen der Europäischen Union, Luxemburg (2006)
86. FMVSS 111 National Highway Traffic Safety Administration: FMVSS 111—Rearview Mirrors. U.S. Department of Transportation, Washington (1999)
87. Europäische Kommission: Analyse der Optionen zur Verringerung der Treibhausgasemissionen um mehr als 20 % und Bewertung des Risikos der Verlagerung von CO_2-Emissionen. Amt für Veröffentlichungen der Europäischen Union, Brüssel (2010)
88. Europäische Kommission: Fahrplan für den Übergang zu einer wettbewerbsfähigen CO_2-armen Wirtschaft bis 2050. Amt für Veröffentlichungen der Europäischen Union, Brüssel (2011)
89. R 96/53/EG Europäisches Parlament und Rat: Richtlinie des Rates … zur Festlegung der höchstzulässigen Abmessungen für bestimmte Straßenfahrzeuge im innerstaatlichen und grenzüberschreitenden Verkehr in der Gemeinschaft sowie zur Festlegung der höchstzulässigen Gewichte im grenzüberschreitenden Verkehr. Amt für Veröffentlichungen der Europäischen Union, Luxemburg (1996)
90. ISO 16505 International Organization for Standardization: Road vehicles—Ergonomic and performance aspects of Camera-Monitor Systems—Requirements and test procedures. International Standard, International Organization for Standardization, Geneva (2015)
91. SAE J941 Society of Automotive Engineers: Motor Vehicle Drivers' Eye Locations. Surface Vehicle Recommended Practice, SAE International, Warrendale (2010)

Part IV
CMS Tests and Concepts for Passenger Cars and for Commercial Vehicles

Camera-Monitor Systems as a Replacement for Exterior Mirrors in Cars and Trucks

Eike Andreas Schmidt, Heike Hoffmann, Rainer Krautscheid, Maxim Bierbach, Alexander Frey, Jost Gail and Christine Lotz-Keens

Abstract Camera-monitor systems (CMS) can be used in motor vehicles to display the driver's rear view on a monitor mounted inside the vehicle. This also offers the possibility of replacing conventional exterior mirrors with suitable CMS and thereby implementing new design concepts with aerodynamic advantages. However, as exterior mirrors are safety-relevant vehicle parts for securing the driver's indirect rear view, the question arises whether CMS can provide an equivalent substitute for mirrors. In the scope of this study, CMS and conventional exterior mirrors were compared and assessed in test drives and static tests under different external conditions. On the one hand, the examination of technical aspects, and on the other hand, issues pertaining to the design of the human-machine interaction, were the objects of the study. Two vehicles were available for the trials with passenger vehicles: A vehicle, manufactured in small series, which is already equipped with CMS as sole replacement for the exterior mirrors, as well as a compact class vehicle which had a CMS retrofitted by the car manufacturer in addition to conventionally used exterior mirrors. The latter could be covered exclusively for trips with CMS. A tractor unit with semitrailer was available for the truck trials. The driver's cabin was equipped with a CMS system developed by the vehicle manufacturer. In general, it was shown that it is possible to display the indirect rear view sufficiently for the driver, both for cars and trucks, using CMS which meet specific quality criteria. Depending on the design, it is even possible to receive more information about the rear space from a CMS than is possible with mirror systems. It was also shown that the change from mirrors to CMS requires a certain period of familiarisation. However, this period is relatively short and does not necessarily result in safety-critical situations.

Keywords Camera monitor systems CMS · Exterior rear-view mirrors · Human-machine interaction · Technical aspects of CMS · Rear field of vision ·

E.A. Schmidt · H. Hoffmann · R. Krautscheid · M. Bierbach (✉) · A. Frey · J. Gail · C. Lotz-Keens
Bundesanstalt für Straßenwesen (BASt), Federal Highway Research Institute, Bergisch Gladbach, Germany
e-mail: bierbach@bast.de

A. Terzis (ed.), *Handbook of Camera Monitor Systems*,
Augmented Vision and Reality 5, DOI 10.1007/978-3-319-29611-1_12

Comparison of mirror and CMS properties · Image reproduction · Last safe gap methode · Glance behaviour · Eye-tracking

1 Introduction

Camera-monitor systems (CMS) can be used in motor vehicles to display the driver's rear view on a monitor mounted inside the vehicle. This also offers the possibility of replacing conventional exterior mirrors with suitable CMS and thereby implementing new design concepts with aerodynamic advantages. However, as exterior mirrors are safety-relevant vehicle parts for securing the driver's indirect rear view (requirements specified in UN Regulation No. 46), the question arises whether CMS can provide an equivalent substitute for mirrors. Therefore, the Federal Highway Research Institute (BASt) was commissioned by the Federal Ministry of Transport and Digital Infrastructure (BMVI) to carry out corresponding investigations, in which CMS and mirrors are evaluated comparatively. Tests with vehicles equipped with CMS, mirrors or both, were conducted for this purpose. On the one hand, the examination of technical aspects, and on the other hand, issues pertaining to the design of the human-machine interaction, were the objects of the study.

2 Literature Analysis

2.1 Technical Background

According to Regulation No. 46 (R 46) of the United Nations Economic Commission for Europe (UNECE), "Uniform provisions concerning the approval of devices for indirect vision and of motor vehicles with regard to the installation of

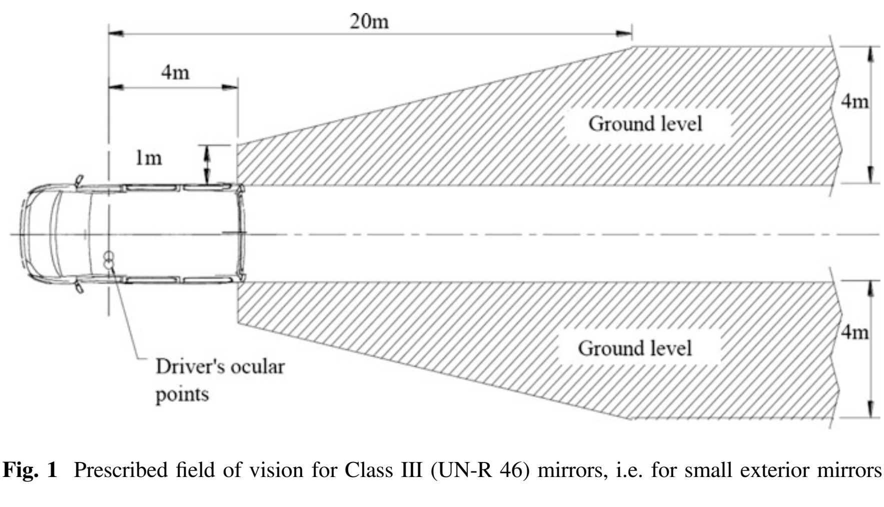


Fig. 1 Prescribed field of vision for Class III (UN-R 46) mirrors, i.e. for small exterior mirrors

these devices" [32] different mirrors are classified into groups according to their purpose. It is stipulated which of these mirrors must be present in the different vehicle classes.

The UN-R 46 defines exterior mirrors as mirrors mounted on the external surface of a vehicle intended to give the driver a clear view to the rear, side or front of the vehicle within a clearly defined field of vision. Figure 1 shows an example of the prescribed field of vision for indirect vision for cars.

According to UN-R 46, a CMS, i.e. "camera-monitor device for indirect vision", is defined as a device which represents the field of vision obtained by means of a camera-monitor combination to the driver. Camera-monitor systems are used in vehicles in order to provide the driver with information on a specific field of vision (usually the rear view). However, at present it is not permitted to use CMS as a replacement for exterior mirrors. CMS may only be used as an additional source of information for the driver.

The International Organization for Standardization (ISO) published a standard with the subject of CMS (ISO standard 16505 "Road vehicles—Ergonomic and performance aspects of Camera-Monitor Systems—Requirements and test procedures") [22]. The standard deals with the requirements and test procedures for CMS in road vehicles.

2.2 Human-Machine Interaction

The first studies on human-machine interaction with regard to CMS were conducted already in 2002. It was concluded that CMS could offer many potential advantages to the driver [11] . However, with a CMS the driver is not able to change the indirect vision by moving his/her head, as it is possible with a mirror. In addition, the position of the displayed image is completely different. The position of the exterior mirror is outside of the cabin; whereas the position of the monitor would be inside the cabin and closer to the driver.

2.2.1 Glance Behaviour in Real Traffic Situations

According to the regulations of the **A**lliance of **A**utomobile **M**anufacturers [1] for the design of human-machine interfaces in vehicles, glances onto the visual display must not be longer than 2 s. A complete secondary task which is performed during driving should not exceed 20 s. The maximum time of visual orientation towards a secondary task, which is usually accepted by the driver, is about 1.5 s. Regardless of whether the searched information, was mentally processed, the driver would usually return his/her gaze to the driving task [27].

If the complexity of the driving task increases, the drivers' frequency of single glances at the display instead of glance duration increases. According to Zwahlen et al. [37], tasks which require up to three glances and mean glance durations of

1.2 s are acceptable. Tasks which require three to four glances and mean glance durations of 1.2 to 2 s are at the threshold of acceptance. If more than four glances and mean glance durations of more than 2 s are required, the task is considered safety-critical. Thus, the mean glance duration and glance frequency are valid values for the detection of critical gaze patterns.

2.2.2 Glance Behaviour During Lane-Changing

Eye movements and glance movements either occur as an adaptation to body and head movements in order to ensure the continued fixation on a target, or as micro movements of the eyes during fixation in order to maintain sensitivity towards a continuous visual stimulus. These micro movements occur during every fixation. Also eye movements particularly occur as a response to the focus on an interesting target. Here, the continuous eye movement allows for acuteness of vision, because foveal perception is only available in a deviation of approximately 1°−1.5° from the visual focus.

In addition to foveal perception, other factors play a role while driving a motor vehicle. For example, peripheral objects are perceived by motion stimuli or contrasts and are classified as conspicuous stimuli [28]. However, peripheral objects are perceived with a lower resolution and colour intensity.

In research, glance and head movements often serve as indicators for imminent lane-changing. According to Henning et al. [18], lane-changing determines quick glance alternation between the left exterior mirror and the road. Immediately before the driver initiates lane changing, the driver glances over his/her shoulder before he moves the steering wheel. In a study by Bayerl [3], five subjects performed 650 lane-changes to the left. Regarding to the results, the number of glances into the mirror increased approximately 1−5 s prior to the actual lane-changing (see Fig. 2).

Henning [19] also examined the earliest indicators for lane-changing and found an increase of glances into the left exterior mirror and a corresponding decrease of glances at other objects. In addition to the glance behaviour, the study shows that the

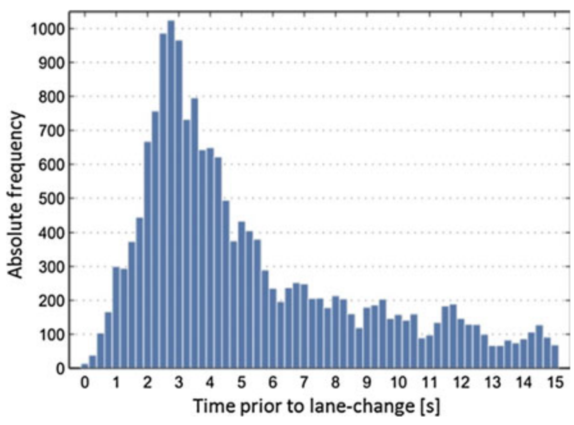


Fig. 2 Temporal distribution and absolute frequency of glances into the *left* exterior mirror within the last 15 s prior to the actual lane-changing. According to Liebner et al. [23]

point in time of crossing the lane, as well as the gradual change of the steering angle, was well predictable. However, it was also found that switching on the indicator or a glance over the shoulder are relatively poor indicators for lane-changing [20].

2.2.3 Distance and Speed Perception in Road Traffic

In research, the “last safe gap” method (see [4, 5, 9, 25, 26, 34]) has proven itself for determining the perception of distance and speed via the rear-view mirror. Hereby, the subjective distance and speed perception, which is difficult to determine by humans, is captured indirectly by means of indication of the last safe moment for lane-changing. The subject sits in a static (semi-dynamic design) or dynamic (dynamic design) test vehicle and, by pressing a key button, indicates the last possible moment for him/her to change a lane in view of an approaching vehicle. The approaching vehicle is only observed via the left exterior mirror. It must be noted that in (semi) dynamic tests, it is difficult to differentiate the perception of distance from the perception of speed, because the perception of different speeds of an approaching vehicle occurs as a result of a simultaneous change in distance. The ‘last safe gap’ method [2] complies with this assumption.

Due to the curvature of the mirror surface, objects are perceived smaller, which complicates the estimation of speed of approaching vehicles [25]. In the studies by Mortimer and Jørgensen [26] and Mortimer [25], the subjects were also allowed to use the interior mirrors for estimations. Here, it was shown that even by using different types of mirrors (planar; 47 and 29 in. spherical), there were no differences in the estimation of the last safe gap. Accordingly, drivers rely more on the interior mirror when estimating the distance and speed of other road users. If subjects estimated without using the interior mirror, the distance to other vehicles was overestimated. This overestimation of distance was greater in mirrors with a smaller radius of curvature. Previous experience with non planar mirrors can partially reduce the overestimation. De VOS [7] examined the compensation ability due to previous experience for estimation of distances. Regarding to the results, subjects who were familiar with spherical mirrors compensated the reduction of size of the objects by selecting significantly larger gaps for lane-changing than subjects who had a planar or partially aspherical mirror on their vehicles. Flannagan et al. [14] also showed that increasing the drivers’ experience with the mirror type, the estimations became more accurate.

Flannagan and Mefford [12] examined the influence of the displayed object size on distance perception in camera-based monitors in real-traffic conditions. The subjects observed a vehicle, which was preparing to overtake, via the camera-based monitor and responded at which point in time they would still pull out in front of the vehicle. Magnification of the image section by a factor of 1.5 resulted in a significant underestimation of the distance. Whereas, the last possible moment for changing a lane was at a distance of 37.5 m by an image display of 1:1 in relation to the own vehicle, at an image display of 1:2, the distance was only 27.5 m.

2.2.4 Distance and Speed Perception for Exterior Mirrors and Displays

Distance and speed estimations in road traffic show a significant discrepancy between subjectively perceived and objectively measured values. Hereby, distances in the mirror tend to be more overestimated [7, 10]. When estimating speeds, an effect of dimension is seen. High speeds are relatively underestimated [30] and slower speeds are overestimated [17]. Likely reasons for different perception of distance and speed using exterior mirrors or camera-monitor systems are due to the limited availability of depth information provided by the CMS. An essential criterion of depth is a result of the eye movement: Both the convergence of the eyes and the curvature of the lens play an important role in the perception of distance and speed of objects: As distant objects are displayed smaller on the eye's retina, humans perceive the movement of distant objects less than those of close objects, which appear larger on the retina [13]. With more experience, the brain learns to compensate for these differences. The ability to perceive depth is limited for indirect vision, i.e. a glance into the rear-view mirror during driving. The stronger the curvature of the mirror, the higher the compensation performance of the brain. However, all traffic objects are appear to be on one level on a monitor, irrespective of their actual distance—therefore the image appears flat and remains unchanged in relation to the head and eye movements of the driver.

There is evidence to suggest that depth criteria such as accommodation play a relatively minor role for the driver compared to monocular depth criteria (e.g. relative size, light and shadow effect of an image). In view of the above this is interesting considering the described development towards camera-based monitors as an alternative to exterior mirrors, as CMS eliminate or distort oculomotor, stereoscopic and motion-induced depth criteria due to the two-dimensional representation. However, monitors are very good at reflecting monocular depth criteria such as for example the height in the field of vision [15]. When the driver looks at the monitor, the information from the convergence of the eyes, the accommodation, the motion parallax and the information from the retinal disparity would serve to indicate that the objects in traffic are located at a certain distance to the driver. Consequently, an impression of depth is created as a result of additional information such as monocular depth criteria, light and shadow effects and experience with distance distortion [13].

3 Technical Aspects

In this study a CMS is compared to a conventional exterior mirror with regard to the technical and road-safety relevant characteristics. In a CMS, the exterior mirrors are each replaced by a small camera. The image recorded by the camera is displayed in the interior on monitors, for the right and left view, in order to present the driver with a rear view.

In this study the technical characteristics and requirements for such a CMS are examined and evaluated. Hereby, the focus is on the comparison to conventional exterior mirrors. Potential problematic conditions when using CMS are examined more closely.

3.1 Test Vehicles

3.1.1 Cars

Two vehicles were available for the trials with cars: A vehicle, manufactured in small series, which is already equipped with CMS as sole replacement for the exterior mirrors, as well as a compact class vehicle which had a CMS retrofitted by the car manufacturer in addition to conventionally used exterior mirrors (see Fig. 3). The latter could be covered exclusively for trips with CMS. On the vehicle that was equipped for the examination with cameras, the cameras were positioned on the left and right, below the mirror housings.

The two monitors of the test vehicle could be mounted variably as they could be fixed on each side in the relevant position by means of three brackets. The monitors on the driver side and close to the passenger door were defined as Position 1 (CMS 1) (Fig. 4). Position 2 (CMS 2) (Fig. 5) was integrated close to the steering wheel on the air vent grids. Suction feet were mounted on the windscreen, close to the A-pillar of the vehicle for the integration of Position 3 (CMS 3) (Fig. 6). Figure 7 shows the locations of the monitors in a schematic overview. During the study, a monitor was always attached to the driver side and passenger side, in accordance with a left and right mirror replacement system.

On the monitors used here, one section is represented as a spherical figure and another section as an aspherical figure, in analogy to the mirror. The external aspherical area contained 200 horizontal and 480 vertical pixel. Whereas the spherical area contained 600 horizontal and 480 vertical pixel (KMS Type DASP quoted from TÜV Test Report, [31]). The transition from the main surface to the aspherical area was clearly marked (Fig. 8). The camera shot 23.5 images in a dark environment and 33.3 images/s in a light environment. In comparison, from an

Fig. 3 *Left* exterior mirror with camera mounted to the mirror housing

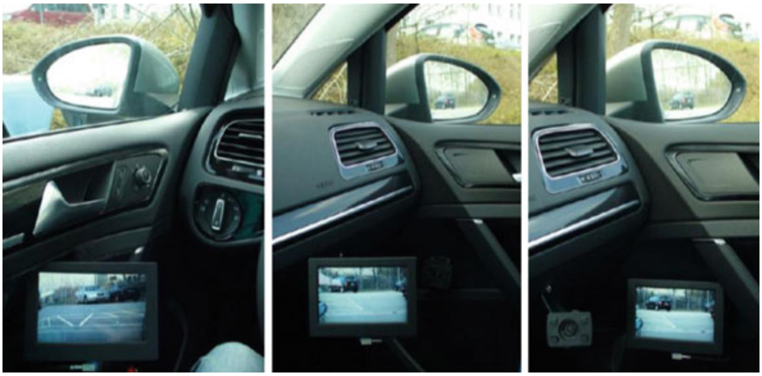


Fig. 4 Integration of the monitors in the driver and passenger door (Position CMS 1)

Fig. 5 Integration of the monitors on the air vent grids (Position CMS 2)

Fig. 6 Integration of the monitors next to the A-pillar (Position CMS 3)

Fig. 7 Schematic overview of the location of the positions CMS 1, CMS 2 and CMS 3

Fig. 8 Right-hand monitor with aspherical section

image frequency of approximately 14 to 16 images/s, the human brain perceives consecutive images as a moving scene [8].

A driver-side partially aspherical exterior mirror with a curvature radius of 1260 mm (in the spherical section) and a passenger-side cylindrical convex exterior mirror were used as exterior mirrors. There was also a conventional interior mirror.

3.1.2 Trucks

A tractor unit with semitrailer was available for the truck trials. The driver's cabin was equipped with a CMS system developed by the vehicle manufacturer. The cameras were fixed above the main exterior mirrors on the driver and passenger sides. In addition, the vehicle had conventional exterior mirrors, which could also be folded away for trips with the CMS in such a way that the driver was no longer able to use them (see Fig. 9).

The system was based on an embedded FPGA (Field Programmable Gate Array) system (processor with higher graphics operations) and adapted firmware. The displays and cameras consisted of components developed by the manufacturer (two cameras with a resolution of 1.3 mega pixels, manufacturer APTINA; two 12.3 in. monitors with a resolution of 1440 × 540 pixels, visible area 295 mm × 113 mm).

Fig. 9 Exterior mirrors with cameras and monitor integration next to the A-pillar on a truck

3.2 Test Concept

The direct comparison of the CMS and the mirror is an important factor for the evaluation. To this end, test drives under various conditions as well as static tests were performed [24]. The static and dynamic tests were recorded as video or photo material for analysis. In order to best document the rear view, an additional camera was mounted and aligned in such a way, that both the left rear-view mirror and the left monitor were visible in the recording (see Fig. 10). The necessity of the documentation required to reproduce the mirror image or the monitor image, which are actually created in the driver's eye, as an "image of an image" in the report. Therefore, the impression in the test driver's eye is decisive for the assessment of the CMS and mirror. Accordingly, the "images" of the mirror images and monitor images in this report show, in parts, the described effects equally clear, however, in some parts also to a limited extent.

Generally, for static tests a colour and a grey level scheme were used as a background, in order to ensure a good evaluation of the CMS and mirror images. Furthermore, static tests on interference immunity (e.g. with regard to electro-magnetic compatibility (EMC) or glare) were performed.

The CMS was mainly observed for the following technical aspects and situations —always taking differences to the mirrors into account:

- Rear field of vision and, if necessary, restriction of the direct view forward
- General day and night properties
- Image reproduction
- Glare from other head lamps at night
- Reflections and glare (display)
- Adjustability of the camera and display
- Failure safety
- Behaviour under different weather conditions
- Effect of soiling

The technical aspects were examined exclusively in cars, however, the results equally apply to the CMS on the truck.

Fig. 10 Camera mounting and image for testing the technical aspects of cars

3.3 Mirror and CMS Properties

In principle, mirrors and CMS both have specific advantages and disadvantages, which already allow for a comparison based on technical-physical aspects. Table 1 compares these basic properties of mirrors and camera monitor systems, based on

Table 1 Comparison of mirror and CMS properties

Mirror	Camera-monitor system
The law of reflection applies to mirrors. A convex curved mirror provides the viewer with a reduced virtual image of the object. The mirrors can be adjusted to adapt to the user's need. The field of vision can be changed by moving the head	A camera records a constructively specified field of vision, which is displayed on the monitor. Moving one's head does not alter the displayed field of vision. However, setting options for the camera are conceivable. Also, an adjustable design of the monitor is possible, in order to e.g. ensure a orthogonal line of sight. Deviating from the orthogonal angle of view on to the monitor can result in an altered perception of the depicted objects, contrasts, luminance and colours
Mirrors depict the object luminance according to its degree of reflection. The object luminance, which is dependent on the degree of reflection, is perceived by the eye, however, diminished by the transmittance of the side window	The maximum luminance of monitors is limited. Ambient light can diminish the luminance contrast and the colour saturation on the monitor. Also, at night, the monitor has a basic luminance greater than 0 cd/m^2
Light on the mirror, e.g. sunlight or light from other vehicles can result in a physiological glare for the driver. Point light sources, e.g. dipped beam head lamps, are reflected as point light sources in the intact mirror. In mirrors with a quality that is standard in the automotive sector artifacts do not appear	Direct light on the camera can result in artifacts. These artifacts strongly depend on the quality of the camera, especially the lens. Direct light on the monitor can result in a diminishment of the luminance contrast and the colour saturation of the image. Furthermore, direct light reflected on the monitor (simple reflection at the cover glass of the monitor) can cause a physiological glare for the driver. Also, especially at night, the monitor image can reflect in the windows and impair the direct outside view
Mirrors reflect colours very well	For CMS the image fidelity is determined by the optical-electrical-optical transfer function. The colour range of a CMS is limited. Colours can be perceived differently by changing the angle of view on the monitor. Artifacts can affect the depiction and perception of colours
The resolution of mirrors is higher than the possible resolution of the human eye	The resolution of CMS is limited and depends on the quality of the components used
Mirror image changes are reflected in real-time	Camera image changes are depicted with a minimal time-delay

(continued)

Table 1 (continued)

Mirror	Camera-monitor system
The degree of reflection of a mirror can be affected by dirt, condensation, scratches, cracks or rain drops. In addition, viewing the externally mounted mirrors can be affected by dirt, condensation, scratches, cracks or rain drops on the side window	The camera image can be affected strongly by dirt, condensation, scratches or rain drops on the CMS. However, viewing the internally mounted monitor is not affected by the side window
A mirror is always operational	The CMS requires time to boot up
Mirrors can fail in the form of scratches, cracks, broken glass or dirt	CMS can fail in the form of missing images (e.g. due to no power or electro-magnetic interferences) or an image containing artifacts

TÜV Test Report [31]. The table lists artifacts. Artifacts are image errors (in parts the image does not reflect reality correctly).

3.4 Fundamentals of Optical Image Effects

Depending on their quality, cameras in CMS have more or less pronounced optical effects, just like other CCD cameras, which affect the representation of visual range. The following describes a few important effects.

Blooming

When photographing extremely light or reflective motifs with a digital camera, blooming effects can occur which let entire surface areas (which are larger than the light motif) appear in a bright white in such a way that contours can no longer be seen. These effects occur, for example, when photographing very bright cloud formations, reflective glass surfaces or mirrored objects.

The blooming effect is caused by an overload of individual photo cells in a CCD sensor. If no light falls on a photodiode, there is virtually no current and the missing current is interpreted as "black". However, if more light falls on the sensor, there is a higher current, up to a maximum value (white). In case of an "overdose" of light, a single photodiode overloads and emits excess electrons on the neighbouring element. Thus, this also produces a maximum current (for brightest white), although it might be exposed to a lot less light. A range is created, in which all CCD pixel produce a maximum current and thus maximum brightness [33].

Smear effect

In terms of digital cameras, a smear or smear effect describes white stripes in an image which emanate from particularly bright light sources in the image range. The cause for this is an unwanted influence on pixels which are actually in the dark, caused by a row-type reading process by the CCD camera chips. One cause is that adjacent rows are subjected to stray light. Another cause is the fact that adjacent

Fig. 11 Flare effect (*left* small aperture, *right* large aperture)

pixels can also be charged during the transport of the electrical loads which code brightness (translation of German version of Wiki [35]).

Lens flare

Visible reflections and light scattering from back lighting in a lens system is referred to as lens flare effect (or "reflection of lens light"). In photographs, lens flare often manifests as star, ring or circular patterns which reduce the colour contrast of the affected areas. The shape of the depicted reflections is also influenced by the aperture blades used: for example an aperture which consists of six elements creates hexagonal patterns (translation of German version of Wiki [36]).

Star effect caused by aperture

Telected (large *f*-number), the more striking the flare caused by the iris (Fig. 11 left). This effect does not occur when using large apertures (small *f*-number) or a circular pinhole (Fig. 11 right).

3.5 Tests and Results

3.5.1 Rear Field of Vision and Direct View Forward

For the analysis of the field of vision depicted by the CMS, the field of vision was visualised and measured using traffic cones. The opening angle of the field of vision was delineated both for the spherical and the aspherical part. The displayable areas at a distance of 20.85 m behind the camera ($\triangleq$ 20 m behind the driver's eye point) were correlated with the width on the monitor in order to determine the reduction in size of the aspherical part. It is evident that the prescribed field of vision, see Fig. 1 in the introduction, was fully captured and thus the requirements for the indirect view are met. The length ratios displayed in Figs. 12 and 13, result in a horizontal reduction by a factor of 5.3 for the aspherical section of the monitor compared to

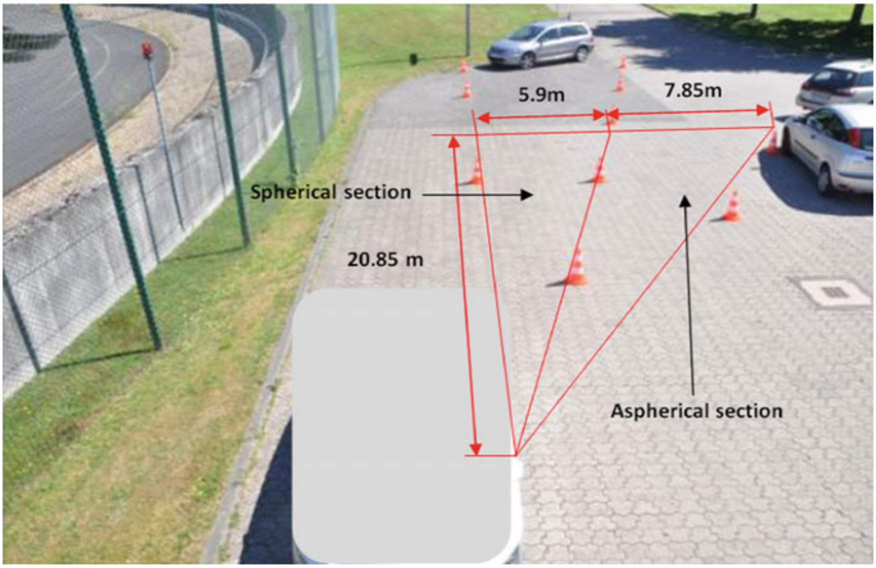


Fig. 12 CMS field of vision

Fig. 13 Monitor view with lateral monitor dimensions

the spherical section. Therefore, as with the aspherical section of the mirror, the so-called blind spot can be reduced. However, an estimation of the distance and speed of objects is more difficult in this aspherical section of the monitor.

No influence on the direct view forward was shown in the study. The monitors in the interior were installed in such a way that they did not impair the direct field of vision.

3.5.2 General Day and Night Properties

The test drives were performed under different conditions. The drives were performed on the following sections and in the following situations:

- Tunnel (jump in ambient brightness)
- Rural road (fast, periodic change from light to dark)
- Motorway
- Low sun
- Uneven road surfaces
- Rain
- Night drive
- Snow

Tunnel (jump in ambient brightness)

How a CMS responds to dynamic changes in the ambient lighting, e.g. when driving through a tunnel, is a subject of the study.

On a bright sunny day the aperture of the camera is only opened slightly before entering a tunnel. When entering the tunnel, the lighting situation changes abruptly (as there is no more sunlight), and the aperture needs to change. On the CMS, the image on the monitor first turns dark, as the camera sensor is underexposed for a moment. Therefore, the image quality decreases in terms of contrast and colour rendering. The camera quickly ($t < 1$ s) adjusts to the darker ambient lighting by opening the aperture. In reverse order, when exiting a dark tunnel and entering bright sunlight, the camera chip is exposed to glare as the opening of the aperture is still too great and needs to be regulated by reducing the size of the opening. The initial overexposure results in a blooming effect, as shown in Fig. 14: whereas the

Fig. 14 Exiting a tunnel during the day

vehicle located to the left behind the test vehicle can be seen in the exterior mirror, on the monitor, it disappears for a short moment in a white field. Also, the white van and the lane markings are more difficult to see on the monitor image in comparison to the mirror, which provides a greater contrast and better colour rendering.

At night, this behaviour is reversed, as a blooming effect occurs when entering the (light) tunnel; when exiting the tunnel the aperture is not opened wide enough yet which results in underexposure. However, these artifacts are not as pronounced at night as the difference in lighting is not as great.

Rural Road (fast, periodic change from light to dark)

A country road through a forest on a sunny day presents a challenge. Much like with the tunnel situation light changes occur here in quick succession between the shadows from the trees and the sunny areas. The aperture needs to change constantly. Until the optimal aperture setting is achieved, over or underexposure occurs resulting in the display of the field of vision becoming impaired. For example, Fig. 15 shows that the quality of the colour rendering and grey shading on the monitor is significantly reduced in the short and bright driving sections due to overexposure.

After entering a wooded area, the generally darker environment and the significantly brighter section in the background (area with no forest) can result in blooming and the associated difficulties in recognition. Figures 16 and 17 show that due to overexposure, a following vehicle can only be seen once it has entered the wooded area. In the mirror, the same vehicle could also be recognised prior entering the dark area.

Motorway

When driving on motorways it is important to detect rear traffic already from a great distance, due to high differences in speed, in order to safely perform driving tasks such as e.g. overtaking.

Fig. 15 Driving on a rural road surrounded by forest

Fig. 16 Entering a forest: The following vehicle is still in the forest-free area

Fig. 17 Entering a forest: The following vehicle is in the wooded area

Early detection requires a good depiction of the traffic. As shown in Fig. 18, it is more difficult to perceive the white vehicle and the lane markings on the monitor compared to the mirror. This is due to insufficient grey graduation of the CMS, as the markings and the vehicle are difficult to distinguish on light asphalt or concrete. It is also evident that the colour rendering in the CMS is significantly worse than in the mirror.

Low Sun

Low sun is considered an extreme lighting situation for driving. When sun rays fall directly onto the camera sensor, the system is subjected to strong glare, which first results in the monitor displaying a complete white image (see Fig. 19) caused by the blooming effect. During the test, no image content could be detected for approximately 2 s. The aperture had only adapted to the sunlight after this period and the content of the rear field of vision reappeared on the monitor, however, with strong artifacts resulting from the remaining blooming effect and smear effect (see Fig. 20).

In this lighting situation the driver does not experience any physiological glare, at any time, from the CMS due to the limited luminance of the monitor. However, exactly this must be noted for the exterior mirror, as due to the law of reflection, the

Fig. 18 Difficult to detect vehicle in the background (white vehicle on beige concrete carriageway)

Fig. 19 Low sun

Fig. 20 Low sun after short adjustment of the camera

Fig. 21 Driving on cobblestones

sun rays are steered into the visual field of the driver and cause glare, even if the sun rays do not shine into the driver's eyes directly. Although this causes severe impairment of visibility, almost always additional information can be obtained from the rear area, as the eye can adapt to this physiological glare.

This shows that both mirrors and CMS have advantages and disadvantages at low sun.

Uneven Road Surfaces

Vibrations that occur for example when driving on cobblestones could result in an error in the image display of CMS. Test drives on uneven road surfaces (here cobblestones in Fig. 21), however, showed no impairment with regard to image sharpness or blurred images. The monitor always displayed a sharp camera image. Only the relative motion between the driver and the monitor or between the driver and the mirror, that occur when driving on uneven road surfaces, resulted in a blurry perception of the rear field of vision. However, this blurriness was equal in mirrors and monitors and can never be fully avoided on uneven roads.

Rain

Rain is a weather condition in which the CMS must function faultlessly. On the one side this refers to the criterion of "waterproof", i.e. the protection against water penetration, and on the other side, to resistance against droplets or formation of water streaking. However, during the corresponding test drives, the need for distinguishing between heavy rain and normal rain became apparent.

Normal rain

During normal rain, rain drops are rarely found on the camera, which generally only influence the image slightly. This is due to the protected installation of the camera.

In the same weather conditions, the view in the mirror is impaired by drops and "water streaking" on the driver's side window. The additional water layer on the relevant surface causes the image to blur and thus results in a worse situation for the assessment of following traffic, as shown in Fig. 22. The monitor view is demonstrably clearer (unaffected) than the mirror view.

Fig. 22 Rear view in rain

Fig. 23 Rear view in heavy rain

Heavy rain

Compared to normal rain, heavy rain causes heavy splashing, which results in more a difficult detection of point light sources in the CMS. This is due to the low brightness of the background (aperture open wide) and the great difference in luminance of the bright head lamps of the following traffic. Point light sources are no longer displayed as such, as shown in Fig. 23, resulting in the overexposed vehicle head lamps in the background melting into one large cone of light on the monitor. However, the overall depiction of the scene was good in the CMS in heavy rain.

As is the case with the CMS, the rear mirror view is heavily impaired by splashing and additionally by rain drops, however, the colour rendering is more realistic due to the better contrast ratio. Therefore, individual vehicles can still be distinguished (see Fig. 23).

Overall, both mirrors and CMS do not provide a good rear view.

Night Drive

At night, just like during the day, the rear view must be reflected in accurate detail in the CMS. Here, it is important that individual road users (vehicles) are

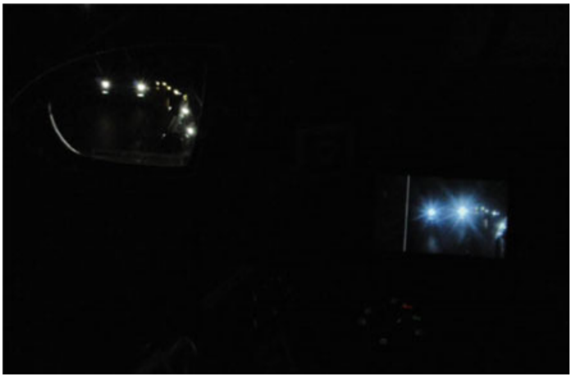


Fig. 24 Driving at night

recognised as such. This means that point light sources of dipped beam or high beam head lamps must be easy to assess (e.g. to distinguish between single and two track vehicles). During a corresponding test drive (see Fig. 24), the individual head lamps of the other vehicles can be recognised well both in the mirror and in the CMS. Only in the CMS, a light flare occurs around the head lamps, however the driver is still able to make out the individual vehicles and to distinguish them.

Additional rain at night causes a merging of the point light sources, in the same way as heavy rain during the day. This makes it a lot more difficult to identify the individual head lamps, which in turn can affect the estimation of speed of the following vehicles. More static tests on glare and point light sources at night are described in section "Camera Glare Caused by a Second Vehicle from the Back with Dipped Beam and High Beam Head Lamps".

Behaviour in Snow and Fog

CMS and mirrors must also be available to drivers in snow and fog. Behaviour in snow and fog was therefore tested as well.

At a low ambient luminance including fogged up side windows and/or droplets on the side mirror, the CMS showed an image that was hardly affected by the weather, compared to conventional mirrors (Fig. 25). The reflection seen on the monitor is caused by the camera flash. The passing vehicle remains more clearly visible than in a mirror.

Fig. 25 Comparison at snow fall and low ambient luminance

Fig. 26 Comparison at snow fall and higher ambient luminance (close vehicle)

Fig. 27 Comparison at snow fall and higher ambient luminance (distant vehicle)

With increased snow fall and higher ambient luminance a vehicle with the dipped headlights turned on, which is close to the camera, merges with the bright background in the CMS. Compared to conventional mirrors the visibility is worse (see Fig. 26).

A comparison image, at almost the same ambient luminance, however, shows less of a difference between the mirror image and the CMS for vehicles that are further away (Fig. 27).

The direct comparison between Figs. 26 and 27 shows the influence of a possibly overexposed camera, due to the badly set head lamp of the passing vehicle, on the display of the scene in the CMS. Therefore, the influence of head lamps causing glare on the image reproduction by the CMS is possibly higher than the influence on the conventional mirrors observed for comparison.

3.5.3 Image Reproduction

For the evaluation of the image reproduction the following technical properties are examined:

- Contrast and brightness
- Colour rendering

Contrast Measurement

Contrast refers to the difference in luminance between bright and dark areas of an image or between to image points. Contrast is defined by the ratio of the maximum and minimum luminance. For statements about the contrast representation of the CMS compared to mirrors, evaluations were performed under different ambient conditions.

The contrast measurement tests were performed in the light hall of BASt and under clear skies during daylight (here for maximum lighting).

In order to reduce the ambient brightness to a minimum, as can be the case in night time road conditions, the light hall was completely darkened. A DIN A0 sized test board, which consisted of alternating white and black squares (see Fig. 28), was positioned at a distance of 1.40 m behind the vehicle rear, in order to perform the measurement of the luminance (in candela per square metre (cd/m^2)) on the individual black or white fields on the monitor and mirror.

The luminance was measured in three different scenarios:

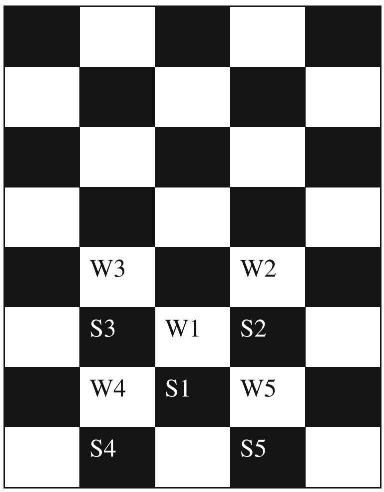
Scenario 1: Turned off vehicle lighting in a dark environment. This scenario represents getting out of an unlit vehicle at night

Scenario 2: Turned on vehicle lighting in a dark environment—represents a normal drive at night

Scenario 3: A clear, cloudless day with bright sunshine. This scenario represents a possible drive on a bright day

For each of the afore mentioned 3 scenarios, the luminance was measured on the test board in five white (W1 to W5) and five black fields (S1 to S5), as marked in Fig. 28. This was done once directly on the board as a reference, once in the mirror and once on the monitor, on the latter, once with maximum and once with minimum

Fig. 28 Test board



brightness setting. An LMT L1009 luminance meter (calibrated in December 2012) was used for the examination.

In scenario 1, a contrast of 1.0 was determined for the CMS and the mirror on the test board, i.e. all fields had the same luminance value "black". The background lighting of the monitor can be determined based on the measurement results from scenario 1 (see Table 2). This has a minimum value of 0.2 cd/m^2 which is significantly lower than the maximum limit of 2 cd/m^2 as required by ISO 16505. The low background lighting avoids reflections in the windows and the monitor lighting virtually prevents a physiological glare for the driver during night drives.

Scenario 2 describes a night drive with low ambient brightness (from the head lamps). The measurement values (see Table 3) clearly show the high contrast on the monitor (14 or 6) compared to the contrast in the mirror (1.6). An internal amplification in the CMS allows for this high contrast value, which displays a contrast that is double as high than in reality (6.4). Even with maximum dimming, the monitor still approximately reproduces the actual contrast. In contrast, in the light conditions of this scenario, the mirror only reflects a quarter of the actual measured contrast. This means, it reduces the contrast.

In scenario 3, the monitor's luminance reaches its limits. The luminance measured on the monitor only has a fractional value of the measurements on the test board. Compared to the CMS, the mirror had a 10 times higher luminance (see Table 4). In the present strong ambient lighting and with the low monitor luminance, the driver's eye needs to adapt more than when looking into the mirror.

In this scenario with maximum ambient lighting, the contrast ratio of the monitor is maximum 50 %, whereas the contrast ratio of the mirror still lies at 80 %. There can be

Table 2 Scenario 1

Scenario 1: Darkness without vehicle lighting				
	Actual value on test board (cd/m^2)	Mirror (cd/m^2)	Monitor with max. brightness (cd/m^2)	Monitor with min. brightness (cd/m^2)
S1	0.1	0.1	0.5	0.2
S2	0.1	0.1	0.5	0.2
S3	0.1	0.1	0.5	0.2
S4	0.1	0.1	0.5	0.2
S5	0.1	0.1	0.5	0.2
Mean of black measuring points	0.1	0.1	0.5	0.2
W1	0.1	0.1	0.5	0.2
W2	0.1	0.1	0.5	0.2
W3	0.1	0.1	0.5	0.2
W4	0.1	0.1	0.5	0.2
W5	0.1	0.1	0.5	0.2
Mean of the white measuring points	0.1	0.1	0.5	0.2
Contrast	1.00	1.00	1.00	1.00

Table 3 Scenario 2

Scenario 2: Darkness with vehicle lighting				
	Actual value on test board (cd/m^2)	Mirror (cd/m^2)	Monitor with max. brightness (cd/m^2)	Monitor with min. brightness (cd/m^2)
S1	0.1	0.2	0.9	0.3
S2	0.1	0.5	0.8	0.3
S3	0.1	0.2	0.7	0.3
S4	0.1	0.2	1	0.4
S5	0.1	0.2	0.7	0.2
Mean of black measuring points	0.1	0.26	0.82	0.3
W1	0.6	0.4	12	1.9
W2	0.8	0.4	6.2	1.1
W3	0.4	0.4	10.6	1.6
W4	0.5	0.5	19.6	3.2
W5	0.9	0.4	9	1.5
Mean of the white measuring points	0.64	0.42	11.48	1.86
Contrast	6.40	1.62	14.00	6.20

Table 4 Scenario 3

Scenario 3: Maximum lighting during the day with sunshine and clear skies				
	Actual value on test board (cd/m^2)	Mirror (cd/m^2)	Monitor with max. brightness (cd/m^2)	Monitor with min. brightness (cd/m^2)
S1	442	189	23.7	12
S2	437	198	30.7	14.6
S3	462	195	23.7	12.4
S4	443	195	21.2	11.1
S5	429	181	31.1	11.6
Mean of black measuring points	442.6	191.6	26.08	12.34
W1	15840	5440	491	63.7
W2	16400	5650	470	63.6
W3	16300	5640	462	63.3
W4	16240	5620	496	65
W5	16080	5570	512	67.2
Mean of the white measuring points	16172	5584	486.2	64.56
Contrast	36.54	29.14	18.64	5.23

several reasons for the mirror contrast ratio not corresponding exactly to the value of test board; due to the law of reflection, the contrast in the mirror should be equal to the test board value. These deviations could be the result of transmission losses when the light passes through the mirror glass, in an imperfect mirror (reflection layer), soiling on the surface as well as slightly deviating measurement points.

In summary, according to this test, different advantageous conditions could be determined for both versions. The CMS is at advantage for night driving with low ambient lighting based on the internal contrast amplification. In comparison, the contrast ratio in bright daylight is better when using the mirror.

Colour Rendering

In road traffic, information is also coded via different colours, e.g. traffic lights or vehicle lighting devices. However, as cameras and monitors only detect or reproduce a certain number of colours, it is important to determine whether these limitations result in deficits.

For this purpose, the colour rendering of the image section displaying pencils in 12 different colour shades was evaluated once by the CMS and once using the mirror (see Fig. 29).

As shown in Figs. 30 and 31, there is hardly any difference with regard to colour rendering between the mirror and monitor. In the mirror image shown here, the colour contrast between the white and orange pencil appears lower than it is in

Fig. 29 Colour range in front of mirror and camera

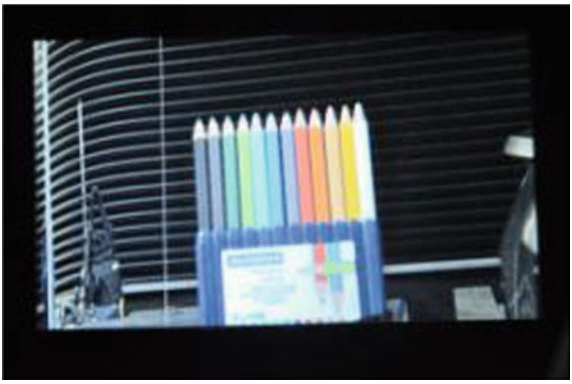


Fig. 30 Monitor colour rendering

Fig. 31 Mirror colour rendering

reality, due to the photography of the situation. On the overall, the colour rendering in the CMS is not as high as that of the mirror, however it is sufficient to render the present colours in a recognisable way. However, this has to do with the good ambient lighting condition, in which all colours are clearly visible. In test drives in less bright conditions, the colour rendering was limited, as shown in Figs. 15 and 18.

3.5.4 Behaviour in Glare

Camera Glare Caused by a Second Vehicle from the Back with Dipped Beam and High Beam Head Lamps

In order to clearly capture vehicles in the rear field of vision during night drives, the head lamps must be recognised as point light sources. To what extent this is the case for CMS and mirrors with high differences in brightness (head lamps versus dark background) was shown in a static investigation with the test vehicle in the completely darkened light hall of BASt and with a second vehicle which represented the following traffic. For this, the latter (Mercedes E-class 1999) was positioned at distances of 50, 25 and 5 m, once with turned on Xenon dipped beam head lamps and once with halogen high beam head lamps (see Fig. 32). This scenario was evaluated both with the CMS and the mirror.

In the CMS representation, the turned on dipped beam head lights at a distance of 50 m, were displayed with artifacts, however, as shown in Fig. 33, point light sources can be clearly recognised. However, a very limited colour rendering of the CMS was shown here under the given lighting conditions. Compared to Fig. 34, in which the same scenario was observed using the mirror, the different coloured traffic signs appearing at the edge of the roadway could be recognised. The dipped beam head light did not cause a physiological glare for the driver in both the CMS or the mirror.

This was followed by the examination of a turned on high beam head light, again at a distance of 50 m. Figure 35 shows that the head lights merge based on the blooming and smear effects which makes the detection of point light sources more

Fig. 32 Test set-up: The following vehicle at a distance of 50 m with turned on dipped beam head lights

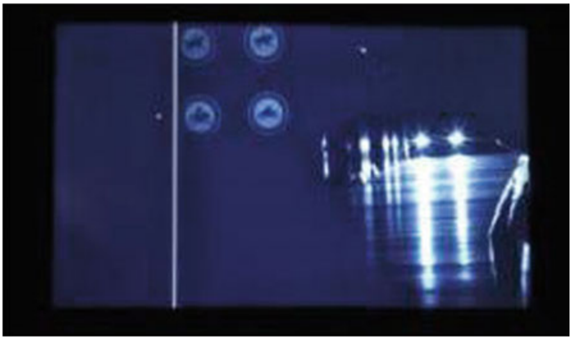


Fig. 33 Monitor at a distance of 50 m and turned on dipped beam head lights

Fig. 34 Mirror at a distance of 50 m and turned on dipped beam head lights

difficult. In the same set-up, both high beam head lights could be differentiated in the mirror (Fig. 36), however, the driver experienced glare and thus a strong impairment of view.

Similar perceptions were determined for both distances of 25 and 5 m, however the intensity of the artifacts decreased: Already at a distance of 25 m, the high beam head lights were displayed clearly as point light sources on the monitor. With regard to the examined conditions it must be stated that the scenario of turned on high beam head lights is unusual in road traffic.

Fig. 35 Monitor at a distance of 50 m and turned on high beam head lights

Fig. 36 Mirror at a distance of 50 m and turned on high beam head light

Camera Glare Caused by Infra-Red

A combination of infra-red lights and special cameras is used in driver-assistance systems such as for example "Night Vision" systems, for the improvement of the recognition of objects and persons during night driving and thus for increasing safety. To this end, infra-red light in the wavelength range of 700–1000 nm is emitted, reflections are detected by an optimised camera and are provided to the driver in an appropriate representation. Whereas infra-red light is not perceived by the human eye, and other road users are not affected [29], the question arises whether the CMS is sensitive for infra-red radiation and to which extent it is potentially impaired by this.

For the examination, an infra-red light emitting remote control was aligned to the camera, by pressing any key, a signal was sent by the infra-red emitter. The infra-red signal was clearly represented on the monitor as a light point.

It was shown that the camera of the CMS is generally sensitive to infra-red light. However, an evaluation with regard to the impact of a driver-assistance system with active infra-red lighting could not be performed, as no suitable system was available. It appears that this type of wave length light could cause glaring—similar to low sun. It still needs to be examined whether glaring through infra-red light can occur and which technical possibilities (e.g. filters) could be used to prevent this.

3.5.5 Reflection on the Display and Screen Glare

Glare Through Reflections on the Display

In some test drives reflections could be observed on the monitor of the CMS. These were so severe that the view out of the window of the driver's door could be seen on the passenger side display due to the reflections. As a result the image content of the monitor could hardly be seen (see Fig. 37). When, for example, turning off to right, this could result in the driver overlooking other road users. Comparable reflections did not occur in mirrors.

Screen Glare

Screen glare means direct light falling on the screen of the display and the consequences thereof. This incidence of light on the screen of the display is usually caused by the sun. As the monitor has a limited maximum luminance, the luminance caused by sunlight on the display can be greater. This results in a reduction of contrast and colour perception for the driver.

The effects caused by screen glare due to direct light incidence are shown in Fig. 38. The image clearly shows the difference between the part of the monitor exposed to sunlight and the shadow of the A-pillar. The colours in the shaded part and the contrast are clearly much stronger. As a consequence, road side vegetation

Fig. 37 Reflections on the passenger side display

Fig. 38 Screen glare

appears more green and differences between bright and dark are more pronounced. In order to minimise the impact of glare due to direct light incidence, the display could be equipped with a shield or set back in the housing.

3.5.6 Adjustability of the Camera and Display

The examined CMS in the small series vehicle could be changed in the two parameters of zoom and brightness. For the image angle, the adjustment was performed automatically, as soon as the vehicle was put into reverse gear. In this case, the view changed to a wide-angle mode across the entire screen, which previously was divided into a spherical and aspherical section (see Fig. 13). The wide-angle mode which is optimised for reverse driving—symbolised by a displayed “R” on the top outer edge—expands the field of vision in such a way that far objects are represented on the monitor, albeit smaller. However, objects which are located closer to the vehicle, which are not in the field of vision during normal mode, are displayed. The zoom factor of both views (for reverse or forward driving) is non-adjustable, this ensures that the statutory prescribed field of vision for normal driving (no gear or forward gear) is covered at all times. An adjustment of the monitor to adapt to the eye positions of the different drivers for a orthogonal view on to the display is desirable, however, this was not possible in the tested case.

The monitor brightness in the examined CMS could be reduced, by holding down the turn on button. If the driver perceived the monitor image as too bright, despite the automatic brightness adjustment, the background lighting of the monitor image could be dimmed with this button. The automatic control adjusted the set brightness point according to the changing ambient conditions. Restarting the vehicle or briefly tapping the power button reset the original brightness level.

If adjustments are necessary must be evaluated carefully. Adjustments allow for the individual adaption to the driver’s requirements. However, this carries the risk that mirrors or CMS are adjusted by drivers in such a way that the indirect view is no longer optimal. The advantage of the CMS is that the optimal (and prescribed) field of vision can always be preset with the default settings. However, depending on the driving situation, a manual adjustment of the image section can be useful. This particularly applies to trucks. (For mirrors, the image can be adjusted both by head movements or adjustment devices). Adjustability of the monitor to the body size of the driver would also be beneficial, in order to ensure the best possible orthogonal view on to the display.

In terms of brightness and contrast, automatic adjustment to the ambient conditions should be standard, additional manual adjustment options were considered as useful.

3.5.7 Failure Safety

During the CMS investigation, occasional outages of the monitor occurred. For approximately 1 s, an error message as shown in Fig. 39, was displayed instead of the camera signal. These outages occurred both on the left and the right side, however never on both sides at the same time. In normal test drives these outages could not be correlated with any specific events or surroundings. The cause could not be established. However, the facts show the importance of failure safety in safety-relevant systems.

Electromagnetic Radiation

The impact of electric and electromagnetic effects or electromagnetic radiation on technical devices are examined for electromagnetic compatibility (EMC). To determine whether unwanted interactions occur in a CMS, tests involving electromagnetic radiation in the wavelength range of data transmission were conducted. As the CMS is a safety-relevant device on a vehicle, functional outages due to electromagnetic radiation are unwanted.

For the assessment of the effects, a test board was positioned behind the test vehicle in the field of vision of the CMS and mirror in order to receive a comparable and defined image. Two radio units of the brand Topcom with a frequency of 446 MHz were used as the source of the electromagnetic radiation, which sent out a "call" close to the camera, monitor and control device. In addition, a mobile phone (Sony Experia Sola) receiving a call, was placed in the same position.

Interferences of the CMS were observed in the form of image errors on the monitor when the radio devices sent out a call in the vicinity of the control device. The image errors are shown in Figs. 40 and 41 and range from flickering including distortion of the image to a complete loss of image where the monitor displays a red X. However, the radiation from the mobile phone did not trigger any image errors.

Image errors only became apparent when the radio device was located closer than 5 cm to the CMS controls while simultaneously sending a call. Even if this combination of events is not very likely and needs not to be considered very often in everyday life, it is shown that it is very important to design the individual

Fig. 39 Image loss

Fig. 40 Error in image caused by electromagnetic radiation

Fig. 41 Loss of image caused by electromagnetic radiation

components of the CMS with appropriate measures that ensure compatibility with electromagnetic influences.

3.5.8 Behaviour in Extreme Cold and Heat

Behaviour in Extreme Cold

The CMS monitors are based on LCD technology (liquid crystal display). This means that they contain liquid crystals in the displays which operate with severe limitations in cold temperatures as the liquid crystals become sluggish due to the cold, which in turn results in the image loading more slowly. To test the CMS in extreme cold, the test vehicle was cooled in a climate chamber to a temperature of −20 °C and conditioned for one night. A test board was positioned as a reference on to a presentation wall, in order to receive a comparable and defined image.

At the beginning of the evaluation, the ignition was started in the climate chamber to start the CMS. Thereafter, a slow movement of the hand was performed at a distance of one metre, thus generating a moving image. The CMS displayed a blurred image of this movement (see Fig. 42), whereas the movement could be observed in the mirror in real-time (see Fig. 43).

In the second part of this test the vehicle engine was started, the mirror heating was activated and the test vehicle was driven out of the hall due to the exhaust emissions. In order to determine the full functionality of the CMS, the time was measured from starting the engine. After driving out of the hall, condensation formed and accumulated both on the camera and the mirrors (see Fig. 44). After two

Fig. 42 Monitor view

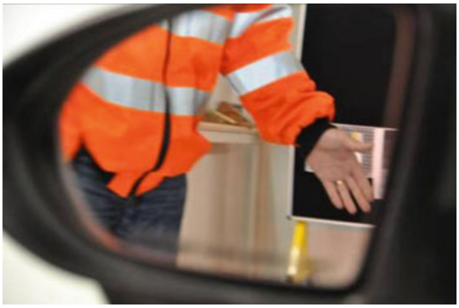


Fig. 43 Mirror view

Fig. 44 Camera and mirror fogged up

minutes the mirror heating had cleared the mirror surface. However, the camera still showed condensation and was not ready for operation (see Fig. 45).

Only after approximately 6 min, rough outlines of the test board, which was positioned behind it, could be recognised (see Fig. 46). At this time the camera was still fogged up and no clear statement could be made about whether the monitor was fully operational and the image was no longer blurred. Even after 13 min the condensation on the camera lens was still present, therefore it was now cleaned manually (wiped clear) with the result of the image sequences displaying in undisturbed quality.

Due to the fogged up camera lens, it cannot be accurately established from which point in time the monitor works without error or delay. The climate scenario used is

Fig. 45 Mirror is clear, camera still shows condensation

Fig. 46 Monitor view after approximately 6 min

rather untypical: The test vehicle was driven out of the climate chamber (−20 °C) on to an open site (15 °C, shade) and humidity precipitated. Such a rapid climate change might occur in cold regions with road tunnels (e.g. the Alps).

However, the test showed that mirrors allow for a complete rear view after two minutes and the CMS camera was still fogged up after 13 min. A heated CMS, in which both the camera and the monitor are headed, could alleviate this problem. This type of heating could significantly reduce the time up to full operational readiness and thus increase road safety.

Behaviour in Extreme Heat

On the sensor chip in cameras, photodiodes convert light into electric power. The heat alone, together with the base voltage in the sensor, discharges electrons. This creates a so-called basic noise of the sensor chips. During normal operation this noise floor is masked by the high number of electron discharges based on the exposure. However, if the sensor heats up strongly, the colour noise can be amplified and impact on the monitor image. The behaviour of the CMS in heat was examined in a test, more exactly the impact of heat on the camera. For this, the test vehicle was placed in the test hall with the CMS turned on. As a comparison motif, a test board was placed behind the vehicle. The camera was heated with a hot air blower on the driver's side (Fig. 47).

Fig. 47 Test set-up for behaviour in extreme heat

At the beginning of the test, the monitor image was viewed, photographed, the temperature on the camera housing measured (20 °C) and thereafter the hot air blower turned on. The distance between the heat source and the camera was reduced under continuous temperature control until the temperature of the housing reached a stable level of 83 °C. To ensure that the camera chip inside the housing reached this temperature, the test conditions were maintained for 15 min. Finally, a re-evaluation and photographic documentation of the image reproduction was conducted. The result showed that after heating no visual changes could be recorded (see Figs. 48 and 49).

Fig. 48 Monitor image before heating

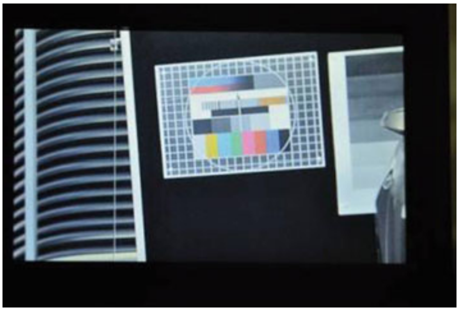
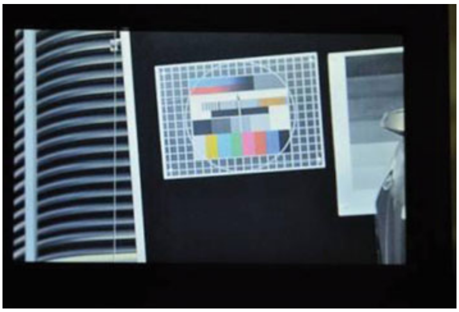


Fig. 49 Monitor image after heating



3.5.9 Effects of Soiling

CMS cameras mounted on the outside of the vehicle as well as mirrors are generally subject to soiling. The type of soiling can have different causes: e.g. pollens, dust, dirt in water or salt. The effects of soiling on mirrors and cameras were analysed in a test series. For this, a test board was placed behind the CMS vehicle and assessed using mirrors and CMS, see Figs. 50 and 51. First, a defined dirt film was applied on the camera and exterior mirror on the driver's side. To create the dirt film, the camera lens or mirror was sprayed with one spray of a saturated salt solution. This was left to dry completely. After assessment and photographic documentation the next step of application was performed. In each of the three soiling stages, the image quality of the CMS must be classified as better, compared to the mirror, as

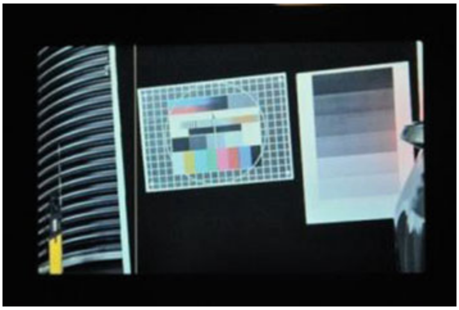


Fig. 50 Monitor image without soiling

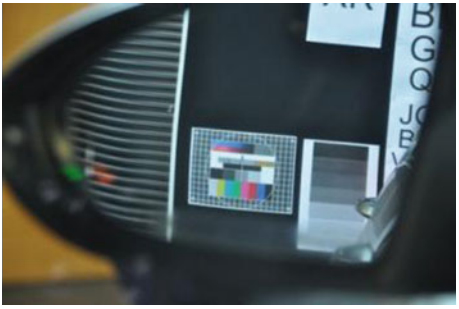


Fig. 51 Mirror image without soiling

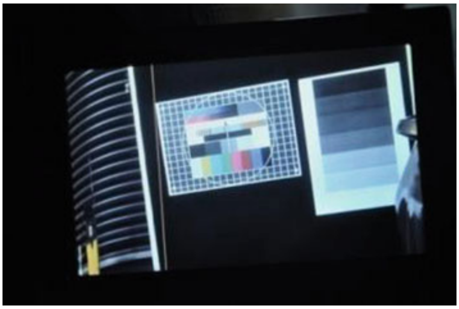


Fig. 52 Monitor image with soiling, step 1

Fig. 53 Mirror image, soiling, step 1

Fig. 54 Monitor image with soiling, step 2

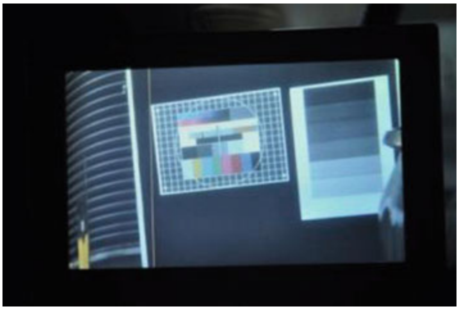


Fig. 55 Mirror image with soiling, step 2

Fig. 56 Monitor image with soiling, step 3

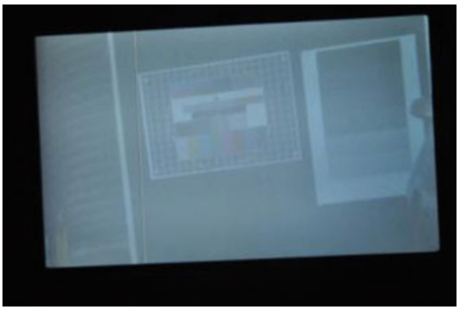


Fig. 57 Mirror image with soiling, step 3

shown in Figs. 52, 53, 54, 55, 56 and 57. The contrast representation and colour representation is higher in the CMS.

4 Aspects of Human-Machine Interaction

Two studies were performed in order to investigate aspects of human-machine interaction during the use of the CMS/exterior mirrors, one study each for the vehicle classes cars and trucks. The underlying research question was, whether the replacement of exterior mirrors with CMS is possible from a HMI perspective. The car study focussed on the estimation of distance, analysis of glance behaviour, and subjective assessment using acceptance criteria. The truck study used estimation of distance and subjective evaluation.

4.1 Car Study

The car study approached the question whether the driver-vehicle interaction changes in case of replacing mirrors with a camera-monitor system [6]. The focus of interest, on the one hand, was on the visual support which the driver received by the system during driving manoeuvres which demand for watching the rear traffic (e.g. pulling out for overtaking). For this, subjective measures (questionnaire) and objective measures (glance behaviour) were taken into account. On the other hand, it was of interest whether changes in distance and speed perception of an approaching vehicle may occur, if the driver solely uses a camera-based monitor. In this case, the assessment of the CMS had to be performed in comparison with conventional exterior mirrors. Furthermore, the question arose whether the driver's ability to appropriately use such a system in road traffic benefits from the driver's previous experience with such a system. Therefore, the sample of the study was divided into two driver groups: (a) *Experts,* who were allowed to exercise the use of the CMS in real traffic before they performed the tests, and (b) *Novices,* who were unfamiliar with the CMS. It was hypothesized that existing experiences affect the driver's ability in such a way that distances are estimated more reasonably (compensation) because of the skills acquired in the learning phase (see, for mirrors: [7, 14]). Accordingly, it was expected that the quality of distance estimation via the CMS converges to the reference values attained via the exterior mirrors over time. Two experiments were performed in the car study. Experiment I focused on distance and speed perception. Experiment II investigated glance behaviour during real drives.

4.1.1 Sample

A total of 42 subjects took part in the study. The total sample was comprised of 18 female subjects and 24 male subjects. The average age of the total sample was 47.8 years with a standard error (*SE*) of 2.7. The youngest subject was 25 years old and the oldest subject was 78 years old.

The sample was first divided into two experimental groups, i.e. an experts group and a novices group. The experts group consisted of eleven subjects, of which 5 subjects were females and 6 subjects were males. The subjects of the experts group were required to pass through a training phase by making use of the CMS in real traffic over a period of at least 60 min up to 2 days at the maximum. The average age of the experts group was 38.4 years ($SE = 2.7$).

The second group was made up of novices. This group consisted of 31 external subjects who did not have any previous experience with the CMS. The novices group only received a short oral instruction on how to use the CMS and a pre-test run.

With regard to visual capabilities, 81.8 % of the experts and 87.5 % of the novices stated that they had no impairments. All subjects could be classified as active drivers (average mileage >12.000 km/year).

4.1.2 Test Procedure

The data was collected on five days per week. Each day two subjects performed their tests, one subject at 9.00 am and the other at 1.30 pm. Maximum duration of the tests per subject was three hours. The subjects received a consent form which informed them about the course of the study and data protection rules. Thereafter, the participants received a questionnaire for the collection of demographic data. The procedures of the main tests are described in the following sections.

4.1.3 Experiment I: Distance and Speed Perception

The first experiment examined changes in the driver's estimates of distance to and speed of a vehicle which approached from behind. Data both during using the CMS and during using a conventional exterior mirror was collected. The estimation using a conventional exterior mirror was used as a reference value. It was of particular interest to determine whether the experts group estimated distance and speed significantly different to the novices group.

The "*last safe gap*" method was applied for the estimation of distance and speed. For this, the driver, while sitting in a static test vehicle, observed the approaching object vehicle by using the left monitor of the CMS or the left exterior mirror of the static vehicle, respectively. Neither the interior mirror nor a glance over the shoulder was permitted for the estimation. The subject had to press a button to indicate the last safe gap before he/she would pull out the vehicle for overtaking (see [2]).

Experimental Set-up

The static test vehicle was parked alongside a test lane at the test facility of BASt (see Fig. 58). Markings were affixed on the ground in order to ensure a consistent rear field of vision.

The object vehicle passed the test vehicle in 12 runs per subject. The number of runs resulted from the number of factor combinations: 3 levels of speed of the object vehicle, in randomised order (20, 35, 50 km/h) × 2 types of devices (monitor, exterior mirror) × 2 trials per speed/device combination. The levels of speed were not given to the subject, but were only known to the test managers. As shown in Fig. 58, the lane on which the object vehicle passed the static test vehicle was divided into (a), an acceleration section (length 50 m), and (b), a target speed section (length 100 m). The driver of the object vehicle was instructed to reach the given speed at the end of the acceleration section and to maintain speed when driving in the target speed section. In order to appropriately interpret the time when the button was pressed, light barriers and reflectors were attached to both vehicles. An additional reflector was located at the 50 m mark of the test lane.

The reflector at the end of the acceleration section triggered the light barrier attached to the object vehicle and was used for the control of the speed conditions.

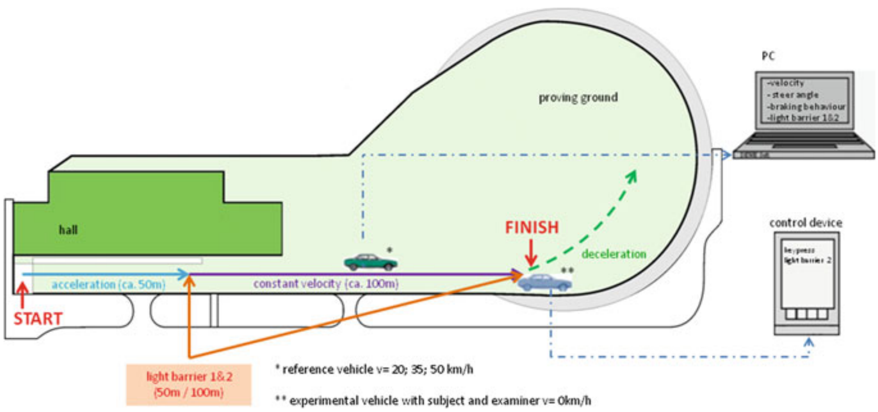


Fig. 58 Schematic test set-up for distance and speed estimation of an approaching object vehicle at the test facilities of BASt

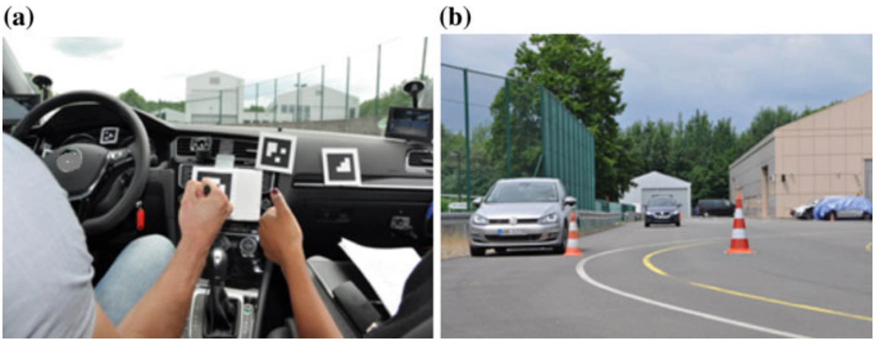


Fig. 59 Static test vehicle at the test facilities of BASt. **a** Inside of the test vehicle: Test participant (*left-hand side*) with button, test manager (*right-hand side*) with observation form. **b** Outside of the test vehicle: Object vehicle approaching from behind at a consistent speed

When the object vehicle passed the test vehicle, the light barrier responded to the reflector on the test vehicle and was triggered a second time. The triggering of the light barrier on the test vehicle served to achieve a real-time synchronisation of the point in time the button was pressed and the point in time the object vehicle passed the test vehicle (*point in time when passing the vehicle—point in time the button is pressed = period used to calculate the distance of the last safe gap).*

Test Procedure

At the beginning, the test manager explained the task to the subject and handed over the button (Fig. 59a). The test manager checked the function of both the light barrier (by means of a reflector film) and the button (response signal) at the control device. Upon arrival of the object vehicle at the start position, the data recording software was launched. The data recording software recorded the following defined channels: speed, angle of the steering wheel, brake actuation and light barrier. The driver of the object vehicle received a list containing the given speeds he/she needed to drive at. A signal (flashing warning lights) was given to the object vehicle driver when to start the next run (Fig. 59b). After six rounds, the object vehicle parked behind the test vehicle and the test manager double checked the correct recording of the data. Thereafter, the object vehicle returned to the start position. The test manager changed the system (mirror, monitor) which was to be used for the next runs. In case the exterior mirror had to be used for estimation, the monitors where turned off manually. If the monitors were used for estimation, the exterior mirrors were covered with adhesive tape. The test manager noted any comments in an observation form as well as the weather conditions, light conditions and subjectively perceived glaring or reflections.

4.1.4 Experiment II: Glance Behaviour During Real Drives

The second experiment examined the driver's glance behaviour in real drives on a motorway (Bundesautobahn). Data were collected both during drives when using the new CMS and drives when using the conventional exterior mirrors. For capturing glance behaviour, an eye-tracking system was used. Particular attention was paid to age-related effects in the novices group.

Experimental Set-up

The drives were conducted between exit *Refrath* and exit *Overath* on the motorway A4. The motorway was travelled on in both directions for this test series. This allowed testing the four considered monitor/mirror positions (CMS 1, CMS 2, CMS 3, exterior mirror, see Fig. 7) on the same route and under similar traffic conditions. The test route one-way from Refrath to Overath was about 14.7 km in length and took a driving time of ca. 12 min.

To exclude sequence effects, the order of the considered monitor/mirror positions were permuted over the subjects. The test event "Overtaking" should occur at least three times and should not exceed five times. The start of the test event "Overtaking" was defined as the moment when the vehicle had completely changed to the left lane. The end of the test event was defined as the moment when the vehicle had completely changed to the right lane.

Test Procedure

After a subject had completed the tests in experiment I, the subject continued with the tests of experiment II. The test manager handed the head unit of the eye-tracking system to the subject. When fastening the head unit on the subject's head, it was ensured that the test participant still had enough space to turn his/her head to the side and glance over his/her shoulder. A calibration device was used in order to align the eye camera to the field camera. The field camera was directed in viewing direction. The eye camera was horizontally and vertically adjusted to the left eye centre of the subject (Fig. 60a).

Thereafter, the line of sight was synchronised to the surroundings (Fig. 60b). For the final check, the subject was asked to glance at the exterior mirrors, the interior mirror and the brackets of the monitors. This procedure was repeated with each change in monitor/mirror position (Fig. 60c).

The subjects drove from BASt facilities to exit Refrath on the motorway A4, followed the motorway towards Olpe until exit Overath, and then returned to exit Refrath. During each drive, the "Overtaking" test events (Fig. 60d) were marked in the logfile by the test manager. The test route was driven several times. After each run the vehicles parked in a parking area at the motorway, where the next monitor/mirror position was set and the subjects completed questionnaires which contained questions about situational awareness. After finalization of all runs the vehicle returned to BASt facilities.

Analysis of the Distance and Speed Perception

The recorded raw data from the eye-tracking system was prepared for statistical calculation by using special analysis programs. Eight subjects had to be excluded

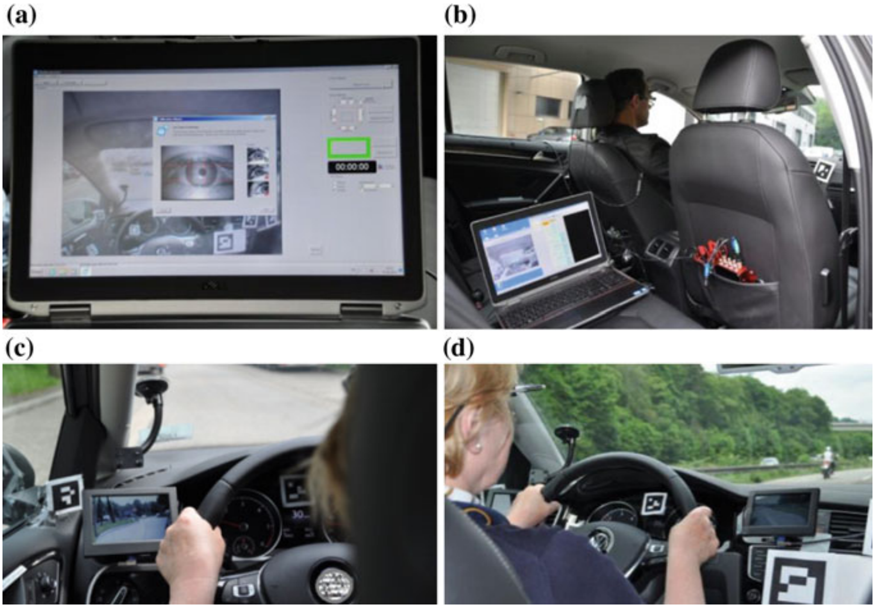


Fig. 60 Measuring glance behaviour by using an eye-tracking system. **a** Calibration of the eye-tracking system. **b** Alignment of the eye camera to the field camera. **c** Monitor mounted on position CMS 2. Test vehicle arriving at the parking area. **d** Test event "Overtaking". Preparing for lane change to the left

from the analysis because of lacking or faulty data. Thus, data records of 34 subjects were included in the calculations.

In order to prepare data for statistical analysis, the signals of the two light barriers and the button needed to be temporally synchronised. This way it was possible to calculate the distance between the object vehicle and the test vehicle at the time the button was pressed.

A multivariate analysis of variance (MANOVA) was calculated in order to determine whether the factors "System type" (2 levels: exterior mirror, CMS) and "Speed" (3 levels: 20, 35, 50 km/h) had an impact on the point in time when the button was pressed. As differences between the measures of the experts group and the novices group were expected, previous experience was defined as sub-subject factor in a further MANOVA with repeated measurements. The *F*-statistic (*Greenhouse-Geisser*) was used in the statistical analysis ($\alpha < 0.05$).

Analysis of the Glance Behaviour

The glance data was validated by checking and manually post-processing pupil detection data for all relevant time lines. The cross-hair marker was readjusted in the centre of the pupil and eye leaps due to incorrect pupil detection were excluded. The images of the field camera were superimposed on the images of the eye camera and examined. Periods which were to be considered for analysis were checked in order to ensure that the glances fell into the defined areas of interest (AOIs). Due to the large volume of data records, only data in time periods needed for analysis were recalibrated. The glance data from the following periods were to be analyzed, because it was known from pre-tests that glances are mainly bound forward when reaching the target lane: a) periods which lasted 10 s until the start of the "Overtaking" event was reached, and b) periods which lasted 10 s until the end of the "Overtaking" event was reached. Selecting a sufficient time period of 10 s ensured that the glances relevant to the changing of lanes could be included in the predefined AOIs of the analysis (see [3]). Thus, the overtaking process was divided into two test events (pulling out, pulling in) for which the following start and end points were defined:

Start "pulling out" = −10 s before the test vehicle had completely changed to the left lane
End "pulling out" = test vehicle had completely changed to the left lane
Start "pulling in" = −10 s before the test vehicle had completely changed to the right lane
End "pulling in" = test vehicle had completely changed to the right lane

The detection of the AOIs and the subsequent calculation of the glances related to the AOIs were coupled to reference points (markers) in the vehicle. The recognition of the markers was performed in an automated process across all glance videos. The glance videos which showed an insufficient detection rate after this automated process were manually processed.

Nineteen subjects were excluded from the analysis, because the data was not usable or the subjects belonged to the experts group. Thus, data records of 24 subjects were included in the calculations. The following glance parameters were considered:

Number of Glances:	Number of glances to the left monitor/left exterior mirror (AOI)
Maximal Glance Duration:	Maximal duration of glances to the left monitor/left exterior mirror (AOI), in seconds

The calculation was performed over all "pulling out" and "pulling in" events for each of the four monitor/mirror positions.

For the further statistical analysis of the data, a univariate analysis of variance (ANOVA) with the factor monitor/mirror positions (AOI) (4 levels: exterior mirrors, CMS1, CMS2, CMS3) was calculated. Here again the F-statistic (*Greenhouse-Geisser*) was used ($\alpha < 0.05$). If the null hypothesis was rejected, the partial η^2 was reported.

4.1.5 Results

Descriptive Statistics

Table 5 Results of the subjects' statements on their traditional use of exterior mirrors in different traffic situations, in percentage (N = 42)

How often do you use the exterior mirror…?	Group*	Never	Rarely	Sometimes	Mostly	(Nearly) always
… for turning	E	.	18.2	9.1	45.5	27.3
	N	12.5	6.3	.	37.5	43.8
… for merging into moving traffic	E	.	9.1	.	27.3	63.6
	N	.	.	.	18.8	81.3
… for monitoring rear traffic	E	.	.	36.4	9.1	54.5
	N	.	.	31.3	31.3	37.5
… before getting out of your car	E	18.2	18.2	.	18.2	45.5
	N	.	6.3	6.3	56.3	31.3

*E = Experts; N = Novices

Table 5 shows the results of the subjects' statements on how they usually use of the exterior mirrors in different traffic scenarios. All subjects stated that they mostly used the exterior mirrors for 'turning', 'merging into moving traffic', 'monitoring rear traffic', and 'before getting out of the vehicle'.

The results of the training phase which the experts group passed in order to gain some experience with the CMS is shown in Table 6. On average the experts used the CMS for 158 min in real traffic. The standard deviation of 96 min indicated a high variance of duration of use. The minimum duration of use was 1 h, the maximum duration was 5.5 h. The average distance driven was 161 km (SD = 135 km). The use of the CMS differed over road types. The percentage of distance driven with the CMS was highest for road type "Motorway" (57 %). The percentages for the other road types were much lower ($\bar{x}(City) = 19\ \%$); $\bar{x}(Ruralroad) = 24\ \%$).

Experiment I—Results of the Distance and Speed Estimation

No significant difference, in terms of distance at the last safe gap, could be demonstrated between the system types CMS and exterior mirror ($F(1, 33) = 3.646$, $p = 0.065$). A significant main effect could be shown for the factor speed of the object vehicle ($F\ (2.66) = 39.752$, $p = 0.000$). The interaction of system type and speed ($F(2,66) = 1.187$, $p = 0.310$) showed no significant effect.

The last safe gap is slightly larger with CMS than with exterior mirrors. The results suggest that subjects would not pull out for overtaking at an earlier point in

Table 6 Duration, distance, and percentage of distance driven by the experts when exercising the CMS during the training phase (N = 11)

	M	SD	Min	Max	Σ
Total duration (min)	158	96	60	330	1740
Distance driven (km)	161	134	32	523	1772
Inner-city road (%)	19	16	5	50	
Rural road (%)	24	14	5	60	
Motorway (%)	57	20	30	85	

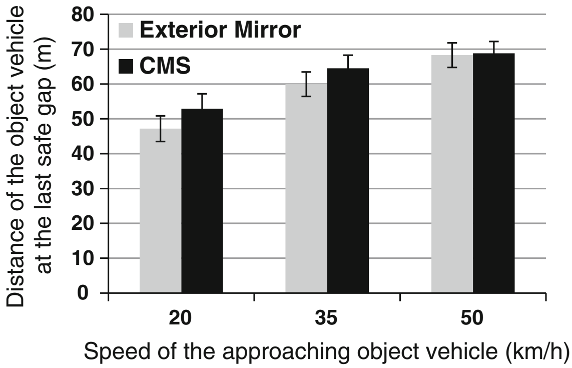


Fig. 61 Last safe gap estimated by the subjects for the two system types. The distance between the moving object vehicle and the static test vehicle was measured at the time the button was pressed ($N = 34$)

time when using the CMS (see Fig. 61). At speeds of 20 km/h, subjects pressed the button at $\bar{x} = 47.2$ m ($SE = 3.7$) when the exterior mirror was used. In contrast, when watching the object vehicle on the monitor, the button was pressed at a distance of $\bar{x} = 59.9$ m ($SE = 4.3$). When the object vehicle approached at a speed of 35 km/h, the distance was $\bar{x} = 59.9$ m ($SE = 3.5$) for the exterior mirror and $\bar{x} = 64.5$ m ($SE = 3.8$) for the CMS. The results showed nearly the same distances in the 50 km/h scenario, i.e. $\bar{x} = 68.3$ m for the exterior mirror and $\bar{x} = 68.8$ m for the CMS. It seems that the distances estimated for the two system types converge with higher speeds.

The results shown in Fig. 61 describe the main effect of the object vehicle's speed on the estimated distance of the last safe gap. Correspondingly, distance increased for higher speeds.

Figure 62 shows the results for the subjects' estimations of distances to a stationary object vehicle (distances: 5, 12.5, 20 m). Significant main effects could be shown for both the factor real distance ($F(2, 14) = 182.3$; $p = 0.000$) and the factor system type ($F(1,39) = 5.203$; $p = 0.028$).

The estimated distance to the stationary object vehicle was smaller when using the CMS than when using the exterior mirror. At the maximum of the given

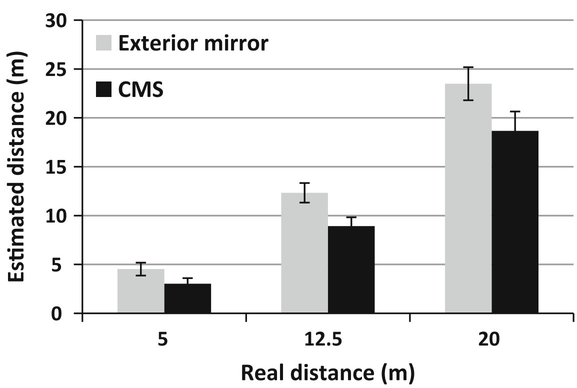


Fig. 62 Estimated distances to a stationary object vehicle using the exterior mirror and the CMS

distances (20 m), the subjects overestimated the distance to the object vehicle when using the exterior mirror.

Experiment II—Results of the Glance Behaviour Tests on Motorway

Due to poor raw data quality the data of only 24 subjects (12 females, 12 males), from a total of 42 subjects, could be used for the analysis of glance behaviour. The average age was 51.6 years (*SD = 16.6).*

Merging into moving traffic

Figures 63 and 64 show the results of glance behaviour during merging into moving traffic. System type showed a significant effect on the number of glances (F (3, 19) = 5.87; p = 0.005) as well as on glance duration of (F(3, 19) = 5.87; p = 0.019).

Overtaking (left lane change)

The mean duration of the overtaking task (time period between pulling out and pulling in again) was 16 s. It did not differ between the two system types (exterior mirrors, CMS). The results of the statistical analysis (Figs. 65 and 66) did not show a significant effect of the system types, neither in terms of glance frequency (F(3, 19) = 2.92; p = 0.06) nor in terms of glance duration (F (3, 19) = 1.65; p = 0.214).

The number of glances was higher with the CMS than with the exterior mirrors, if the monitors of the CMS were located at the A-pillar (CMS 3). This position was close to the position of the left and right exterior mirror. A larger quantity of

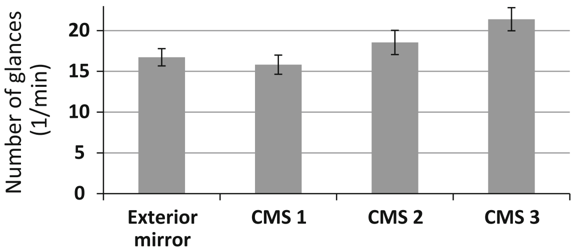


Fig. 63 Number of glances per minute when merging into moving traffic, for different monitor/mirror positions

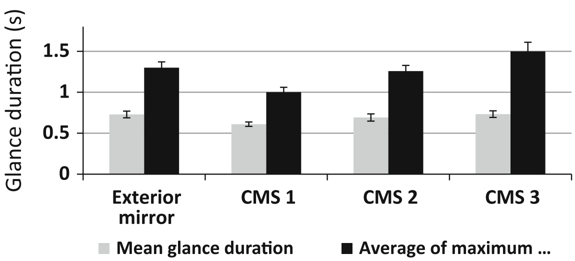


Fig. 64 Duration of glances when merging into moving traffic, for different monitor/mirror positions

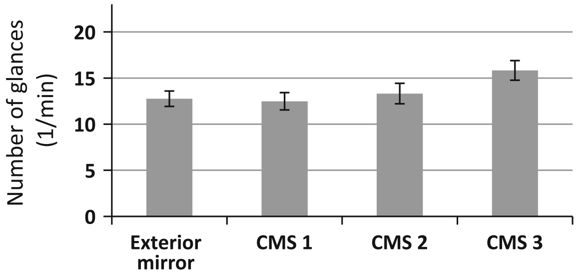


Fig. 65 Number of glances per minute during overtaking, for different monitor/mirror positions

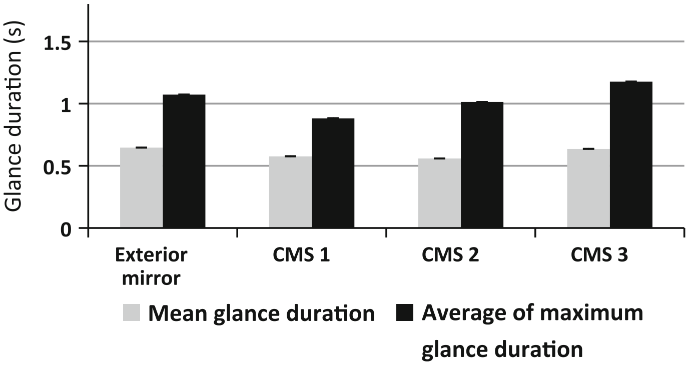


Fig. 66 Duration of glances during overtaking, for different monitor/mirror positions

information on the monitor, due to the wide aspherical area displayed, may be the reason why the subjects needed more glances when using the monitor than when using the exterior mirrors. Glance duration was shortest for the monitor position CMS 1, where both monitors were located at the height of the door panel (see Fig. 7). This means that the monitors which are located below the normal field of view are less taken into account by the subjects. The results show a tendency towards decreased glance duration, if the monitor's position required the subjects to avert the eyes from the vehicle environment and the moving traffic. This suggests that subjects felt more unsafe in this case and, therefore, reduced glance duration.

The analysis of the frequency distribution of the maximum glance duration across all subjects showed that three of 24 subjects exceeded the critical glance duration (2 s) at CMS 2 and CMS 3. At the start of overtaking (left lane change), only two subjects exceeded the critical glance duration during glances at the mirror and at the monitor in position CMS 2. At the end of overtaking (right lane change) the subjects did not exceed the critical glance duration.

Monitor position and acceptance of the CMS

Figure 67 shows the subjects' preferences with regard to monitor position. The most preferred positions were CMS 3 and CMS 2. More than half of the subjects

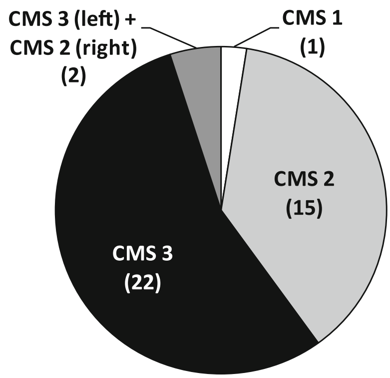


Fig. 67 Monitor position preferences of the subjects (subjective measure)

(22 subjects) preferred CMS 3 and 38 % (15 subjects) preferred CMS 2. Only one subject chose CMS 1. Two subjects gave preference to CMS 3 on the left-hand side in combination with CMS 2 on the right-hand side. The analysis of the acceptance ratings of the CMS showed that acceptance is unrelated to experience gained by the subjects, i.e. the acceptance did not depend on whether and how long the CMS was used by the subjects.

Subjective Assessment of the CMS by the Subjects

The subjective assessment aimed at investigating the subjects' opinion on whether the use of the CMS affected their driver-vehicle interaction. Subjects received questionnaires which were used to collect their statements. The assessment criteria mainly covered the use of the CMS in specific driving manoeuvres and environmental conditions such as parking, lane changing, driving in tunnels, multi-lane roundabout traffic, low sun, night/darkness, contrast/colour. In sum, the subjective assessment showed the following results:

Opinion of the experts group:

Prior to the experiment, the experts group had the opportunity to get used to the CMS. The evaluation of the questionnaires yielded the following results:

- Five of ten experts stated that estimating the distance and speed of vehicles which approach from behind was difficult during changing lanes on the motorway.
- Four of six experts found it difficult to use the CMS during passing through a multi-lane roundabout or driving in curves, as the image representation on the monitor was distorted because of the wide aspherical area displayed.
- Four of nine experts experienced difficulties in estimating the distance during parking.
- Four of seven experts stated that they could not sufficiently recognise the information represented on the monitor when passing through a tunnel, because the vehicle head lights strongly lit up and flashed the environment. In addition, it was hardly possible to recognise anything on the monitor in low background luminance (e.g. lane markings, distance of approaching vehicles from behind).

Opinion of the total sample

The CMS received negative ratings with regard to the criteria of distance/speed estimation and spatial perception. Here, 18 of 20 subjects stated that it was very difficult to estimate different vehicle speeds due to a lack of spatial depth of image representation. Furthermore, 14 of 20 subjects experienced driving in rain as being more disturbing when using the CMS due to the reflections caused by the vehicle lights on the road surface. The majority of the subjects evaluated the CMS positively with regard to the reduced blind spot and the enlarged field of vision to the rear area of the vehicle.

4.2 Truck Study

Following the car CMS study, the CMS properties as well as psychological questions with regard to human-machine interaction (HMI) in trucks were examined.

4.2.1 Sample

A total of 10 male subjects took part in the experiment. All subjects were employees of the BASt. The average age was 51.1 years (*SE = 2.4).* Eight of ten subjects had not driven a truck for an average of 11.4 years. 50 % of the subjects had experience with the camera-monitor system due to their participation in the CMS car study.

Prior to the experiment, all subjects received a demographic questionnaire which contained questions about visual aids, their last consultation to an ophthalmologist, their truck driving experience and routine use of exterior mirrors. All subjects were active car users and hold a class C or class CE driver's licence. With regard to visual function, all participants fulfilled the minimum requirements for visual performance according to Annex 6 of the German Driver Licensing Regulations.

4.2.2 Test Procedure

Before starting the experiment all subjects received the relevant information about the test procedure and data protection regulations. The subjects signed consent forms for participation in the experiment.

Test Procedure

For the evaluation of the CMS, all subjects carried out a test drive at the BASt test facilities as well as in real traffic. The subjects evaluated the CMS based on specified criteria by means of spontaneous statements and questionnaires.

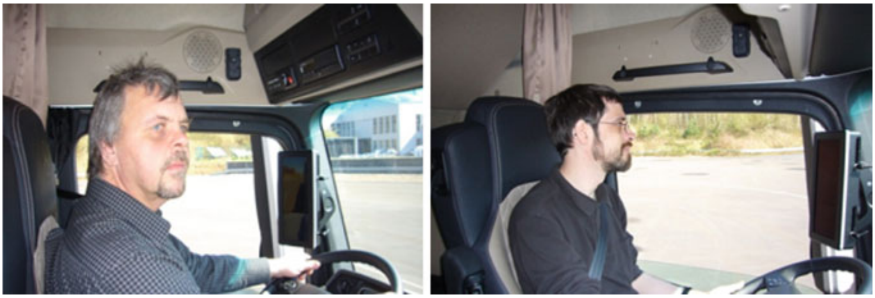


Fig. 68 Subjects during the test drive on the BASt test facility

To get used to both systems, the subjects first completed an exercise drive at the BASt test facilities (see Fig. 68). The exercise drive lasted approximately 20 min and included scenarios such as straight driving, curves and straight reversing.

Before the test drive started, the subjects received explanations about the test procedure, an introduction of the truck operation system as well as information about the camera-monitor system. The total experiment took about 2 h per subject. In nine of ten drives the sun was shining with clear shadow formation. During one of the drives, weather was misty with little sunshine.

4.2.3 Experiment I: Distance Estimation

The distance estimation was performed at the BASt test facility by means of rear approach to two pylons to the right and left of the end of the trailer (see Figs. 69 and 70). A distance of 4 m was selected for the distance estimation.

The pylons had a height of one metre and the distance between both pylons was 3.20 m. For the distance estimation, half of the subjects first started the rear approach to the pylons using the mirrors and then using the CMS; the other half of

Fig. 69 Rear approach to two pylons for distance estimation

Fig. 70 Distance of the trailer to the pylons

the subjects first started with the CMS and then continued with the mirror system. For rear driving using the CMS, the exterior mirrors were folded back.

4.2.4 Experiment II: Commented Drives in Real Traffic

After the exercise drive and experiment I the subjects felt secure enough to drive in real traffic. No subject had aborted a test drive or mentioned that they needed the exterior mirrors as an additional aid for driving.

The first drive in real traffic was performed on the motorway/rural road and served as another exercise for the subjects to get used to the truck and the CMS, because the majority of the subjects had not driven a truck for several years. The first drive was therefore accompanied by a technician, who instructed the subjects in the use of the operating elements and the CMS. This made it easier for the subjects to concentrate on the CMS during the second real drive.

The second drive in real traffic was accompanied by the project manager (a psychologist) who noted any spontaneous statements about the CMS given by the subjects and posed questions to the subjects based on specified criteria (perception of different speeds, driving in roundabouts/urban areas, recognition of distant objects) and documented the answers. At the end of the real drives the subjects received another questionnaire. The total length of route was 57 km, of which 29 km were on the motorway and 28 km on rural roads (see Fig. 71). The route stretched from the Heumarer Mauspfad—direction Cologne/Bonn airport—motorway A 59 direction Bonn/Frankfurt—motorway A 560 direction Frankfurt/Siegburg—rural road 56 direction Much—motorway A 3 direction Cologne—exit Königsforst—rural road towards Bergisch Gladbach-Bensberg.

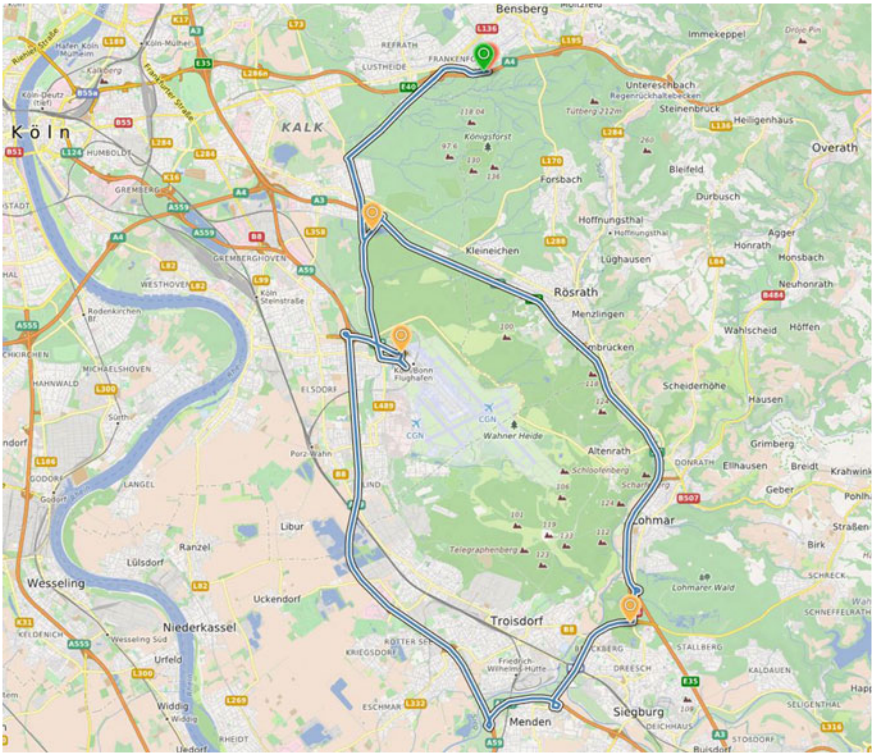


Fig. 71 Test route taken by each subject

4.2.5 Results

The qualitative evaluation of the questionnaires was performed taking the number of positive and negative comments about the CMS and frequency calculations into account. For the analysis of the distance estimation, a paired *t*—test was calculated.

Analysis of Particularities Mentioned by the Subjects

Positive comments

- The system is unfamiliar, however one can get used to it (10 subjects)
- Less soiling (2)
- Disadvantages of the wide angle mirror (distorted image, strong curvature) is compensated by the CMS (1)
- Aerodynamics (1)
- Better direct view out of the windows due to no exterior mirrow mirrors (3)
- Fuel savings (1)
- Front of the trailer clearly visible on the monitor (1)
- No head movement required (1)

Negative comments

- Contrast and colour reproduction too poor; in parts, colours are not realistic (8)
- Shadow formation too strong; road users, objects (kerbs) and distances in the shadow of the trailer are not clearly visible or difficult to estimate (too dark) (7)
- Objects are displayed smaller on the screen (7)
- Display could be larger, especially the right monitor (5)
- Flickering/jittering of the image especially during engine start and turns (3)
- The position of the left monitor is too close, for drivers which are long-sighted and don't wear multifocal glasses; the monitor should be placed closer to the windshield, so that the distance to the eyes is greater (3)
- Position of the right monitor is too far; objects are even more difficult to recognise (3)
- Reduced feeling of safety in comparison to mirrors (3)
- Driving in roundabouts is (rather) difficult (3)
- Dust and finger prints visible and distracting (3)

Subjects' wishes

Manual adjustment for improving contrast, colour and size of objects (2)
Covers on top and at the side of the monitor to prevent glaring (1)

Experiment I—Results of the Distance Estimation

The subjects were asked to reverse up to 4 m to the pylons by using the exterior mirrors and the camera-monitor system. When using the camera-monitor systems, the exterior mirrors were folded back.

Figure 72 demonstrates that short distances (here 4 m) are clearly overestimated when using the exterior mirrors, i.e. a further distance than 4 m is kept to the pylons ($M = 7.5$ m; t-test versus 4: $p < 0.01$). This is not the case for the CMS ($M = 5.5$ m; t-test versus 4: $p = 0.13$). For one subject (subject no. 5) only the estimated value by means of the exterior mirrors could be used for the analysis. The difference between the exterior mirrors and the CMS is not significant, however a clear tendency was shown (paired t-test: $p = 0.062$).

Experiment II—Results of the Drives in real Traffic

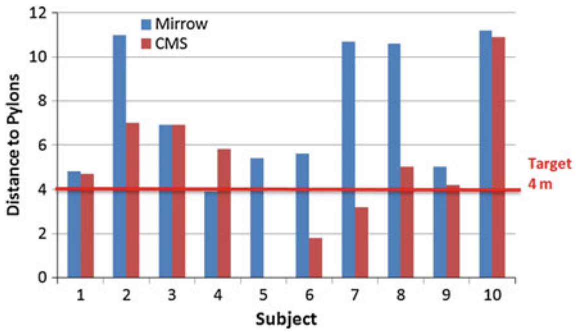


Fig. 72 Estimation of the distance to the pylons

In the following the results of the evaluation criteria of recognisability, colour and image quality, monitor position, driving situation and distance estimation are illustrated graphically.

Recognisability:

Figure 73 shows that the differential speed was better recognised by the subjects than objects which were located further away or than the back of the trailer. This limited recognisability was perceived as disturbing by the subjects (see Fig. 74). One subject gave no statement for the point “Recognisability of the end of the trailer”.

As shown in Fig. 75, nearly 60 % of the subjects mentioned that the recognisability of distant objects was poorer when using the CMS compared to when using the exterior mirrors.

Evaluation of the driving situation:

The majority of the subjects evaluated the use of the CMS in roundabouts or on urban roads as being acceptable. Driving in a roundabout was perceived as difficult by three subjects (see Figs. 76 and 77).

Statement on image quality (brightness/contrast, colour rendering, image sharpness):

Most of the subjects assessed the image quality of the CMS as being poorer than the quality of the exterior mirrors (see Fig. 78).

Information on the display position:

Nearly 40 % of the subjects mentioned that the display position of the monitors was poorer than that of the exterior mirrors (see Fig. 79). 40 % of the subjects would have preferred a left monitor position at a greater distance to the driver;

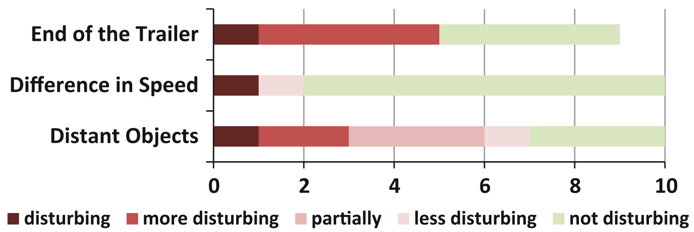


Fig. 73 Estimation of the degree of disturbance on recognisability

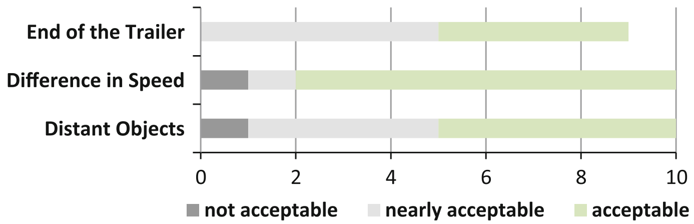


Fig. 74 Level of acceptance of recognisability

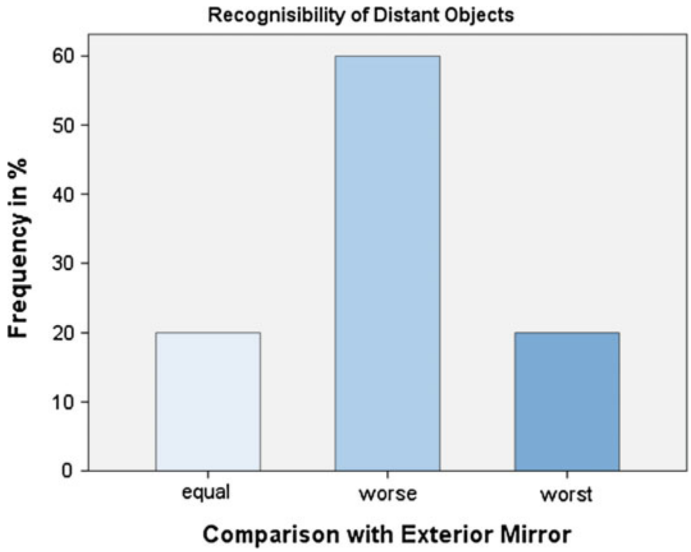


Fig. 75 Comparison between CMS and exterior mirror with regard to recognisability of distant objects

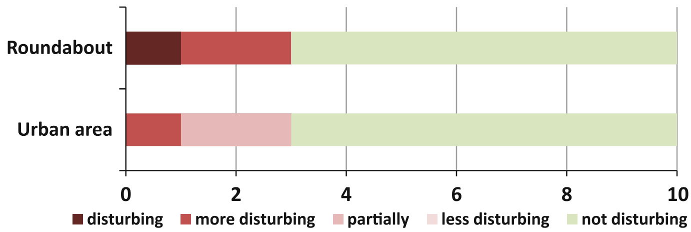


Fig. 76 Evaluation of the degree of disturbance on the driving situation

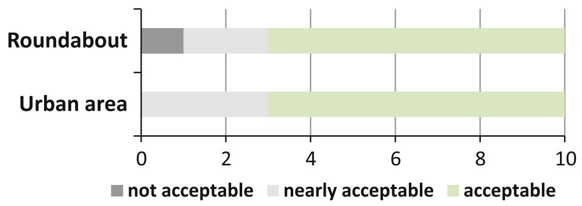


Fig. 77 Degree of acceptance of the driving situation

however, this could be corrected by wearing bifocal spectacles. 30 % of the subjects stated that the right monitor was positioned to far away from the driver. The recognisability of distant objects seemed to decline with the CMS, due to the fact that images of objects were displayed smaller on the monitor than on the mirror.

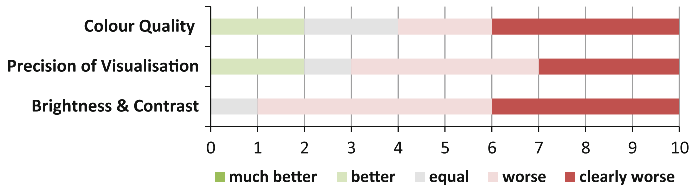


Fig. 78 Comparison between CMS with exterior mirrors with regard to image quality

Fig. 79 Comparison of the CMS to the exterior mirrors with regard to display position

Evaluation of the CMS (support for distance estimation, conveying of a feeling of safety, recommendation):

The green zone in Fig. 80 shows that four of ten subjects stated that the CMS supports the estimation of speed and in total conveys a feeling of safety. Four people at least partially would recommend the CMS in its present form.

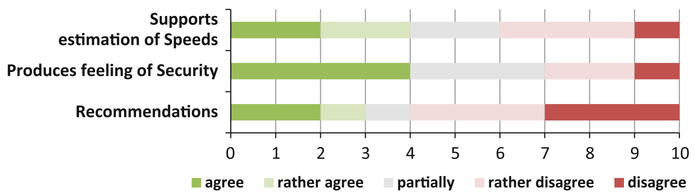


Fig. 80 Assessment of the CMS by the subjects

5 Discussion of the Results

5.1 Technical Aspects

The analysis of the technical properties of the CMS showed advantages and disadvantages of the CMS as compared to conventional exterior mirrors.

With regard to the rear field of vision the CMS covered all required areas and even reduced the blind spot. These are positive features in terms of safety. However, the increased horizontal distortion in the aspherical section of the image displayed on the monitor made it difficult to assess the distance and the speed of the following traffic. It seems to be advisable, in order to shape the transition from conventional mirrors to CMS as smoothly as possible, to design the position of the display of the CMS and its image according to commonly used positions and display formats.

Positions of the cameras and monitors must be chosen in such a way that the all-round view is not impaired and ergonomic requirements are taken into account. It makes sense that the driver does not view the monitor in a skewed manner and that adjustment options allow for an optimal image quality in terms of colour rendering, contrast and luminance.

During operation, a CMS must be able to reliably present information in different and partially changing environmental conditions. Compared to mirrors, there are situations in which the recognisability of the rear field of vision was improved, and in other cases were less good using the CMS.

The prevention of physiological glaring in low sun or turned on high beam head lights of the following traffic, as can occur in a mirror, was assessed as positive. Furthermore, a CMS shows less vulnerability in normal rain with regard to impairment due to water droplets or smearing or soiling.

A dynamic change in lighting, which can occur for example in tunnel entry or exit or in shade from tree-lined roads, is a challenge for CMS. As a result of the interaction of multiple components the optimal display of the CMS in such situations is sometimes only achieved after a certain response time (up to approximately 2 s).

Due to the limited display option of colour nuances and differences in luminance, the study showed situations in which important image details were not recognised or were only difficult to recognise. Here, an improved and more realistic rendering of colour and grey values is desired.

With regard to the displayable contrast, the CMS showed a better reproduction in a dark environment in the tests. In contrast, the mirrors showed a stronger contrast during the day in a brighter environment. Here, the possibility of contrast enhancement for monitors was shown for poor light conditions, and the limitations of the monitor luminance in very bright light conditions.

The reflections caused by light on the covering glass of the monitor were evaluated as negative, as in these situations the image content was not or hardly recognisable. Especially on the passenger side reflections occurred, which could result in overlooking other road users when turning to the right. Remedies are required here, through a different installation location, shielding against sunlight or a reflex-reducing sight protection. In reverse, these measures could contribute to the prevention of the monitor image in the vehicle interior being reflected on other surfaces.

Because a CMS replacing mirrors would be a safety-relevant feature on the vehicle, it is of upmost importance to design it in a way that excludes any outages. It must be ensured that the operational readiness is guaranteed after turning on the system, and that the CMS is operational at all times. This applies to both the power supply and the electrical protection (e.g. fuse). One needs to consider to which extent status monitoring with corresponding signalling for the driver and eventual redundancy needs to be provided. The individual components of the CMS must be designed in an electromagnetic radiation compatible manner.

Within the framework of the experiments, the CMS proved resistant to heat (e.g. by exposure to sun light). However, the sensitivity to cold climate conditions such as thawing, condensation or icing was problematic. Therefore, a CMS heating makes sense.

In order to avoid optical artifacts in image processing, components (lens, camera chip etc.) of quality beeing as high as possible should be used. This also applies to the resolution of the system for a good reproduction of details.

Quick and precise adjustment of the camera to changing ambient lighting and also the aperture is important for optimal functioning. Furthermore, in terms of brightness and contrast an automatic adjustment to the ambient conditions should be standard; additional manual adjustment options were considered as meaningful.

The minimum level of the sampling rate and reproduction frequency for a CMS in order to reproduce timed light signals (e.g. variable message signs or police signs) without loss, was not clarified. It is important the CMS displays the situation without time delay.

If fuel-savings need to be achieved by the CMS, it must be ensured that the energy consumption of the CMS is less than the energy saving due to aerodynamic optimisation.

5.2 Aspects of Human-Machine Interaction

In the present study, aspects of safety-relevant perception using a camera-monitor system as a replacement for exterior mirrors were determined. The subjects' estimation of the last safe gap which they would accept for changing the lane was intended to provide insights on the perception of distances displayed on the monitor.

There was no statistical evidence to confirm the assumption that lane-changes with the CMS, which occur at an earlier point in time than lane changes with the mirror, are assessed as not being safe anymore by the drivers. No significant difference in terms of the used system could be demonstrated. In case of pulling out the car at low speeds, mean distance for the CMS showed a tendency towards an increased safe gap compared to the mean distance for exterior mirror. This result indicates a non-critical misjudgement of speed of and distance to the approaching object vehicle.

Furthermore a highly significant main effect of the object vehicle's speed on the distance at the last safe gap was shown. The distance increased with increasing speed of the object vehicle. The main effect suggests that different speeds can be perceived on the monitors.

As camera-based monitors only provide monocular depth criteria to the driver, they appear to create an impression of depth. Otherwise, the subject always would have to press the button at the same time. Flannagan et al. [13] examined the role of binocular depth information when estimating a relative distance of two vehicles which were viewed in the rear-view mirror. The subject had to provide two estimations, one with one eye and the other with both eyes. It was shown that viewing with both eyes showed no advantage at a distance between 20 m and 80 m. The distances at which the subjects of the present study would still perform lane changes fall in this critical range. Accordingly, based on these results, it must be assumed that the oculomotor, stereoscopic and motion induced depth criteria are not of central significance for drivers in most traffic situations. Consequently, the results of the present study and those of Flannagan et al. [13] suggest that no negative effect for the use of camera-monitor systems is to be expected for the analysed distances.

The overestimation of speed and the underestimation of distance when using the CMS seem to have a positive effect on road safety. As the vehicles are perceived as being closer than they actual are, larger gaps for lane changing were chosen. Possible effects of different traffic densities (congestion, slow moving traffic) could not be conclusively established and need to be considered in further research. If an interior mirror is available to the driver, it can be expected that the additional information available in the interior mirror supports the driver to realistically estimate speed and distance. This may contribute to correct erroneous estimations (see [25, 26]).

The results for the CMS and the exterior mirror show that the distances of the last safe gap converge at speed level 50 km/h. It seems to be of great importance to explore the distance estimation for speeds higher than 50 km/h. The critical question is whether the tendency to underestimate the distance reverses at a certain (high) speed level and turns to an overestimation of the distance. This would have a negative impact on road safety, because vehicles would be perceived as more distant than they actually are. As the recommended speed on German motorways is 130 km/h, a distance and speed perception study according to the applied method is recommended up to this speed level.

There is no statistical evidence for the assumption that distances are estimated more realistically with increasing experience of using the CMS. However, a main effect of experience could be shown irrespective of the implemented system type ($F(1, 21) = 14.673$, $p = 0.001$). The "Experts" pull out later than the "Novices". It is recommended to investigate training effects of using the CMS in further research.

The results of the glance analysis during overtaking (pull out, pull in events) could be summarised as follows: Compared to conventional exterior mirrors, an increased number of glances occurred on CMS 3 only. This position was close to the position of the exterior mirror. A larger quantity of information on the monitor, due to the wide aspherical area displayed, may be the reason why the subjects needed more glances when using the monitor than when using the exterior mirrors. However, the monitors at CMS 1 and CMS 2 did not show that high number of glances, although they displayed the same information. A possible reason for this result can be found when additionally taking the glance duration into account. Maximum glance duration for the CMS 1 and CMS2 (low area of the field of view) was shorter as compared to CMS 3. This can be interpreted as an indicator for lower preference of the CMS 1 and CMS 2 positions due to reduced visual-spatial attention, which is relevant for safe driving. Former studies (e.g. [21]) already verified that visual attention decreases with an increasing distance from the central field of view. It can be concluded that the low preference of CMS 1 and CMS 2 also resulted in a decreased number of glances for CMS 1 and CMS 2.

On the other hand, glance frequency and glance duration for CMS 3 indicate that the monitor located at this position is highly accepted by the drivers. The statements given by the drivers confirm this conclusion.

The truck drives with using the CMS in real traffic proved to be unproblematic for all subjects, i.e. no subject aborted the drive or required the exterior mirrors. Comparing the positive and negative statements, the subjects' majority assessed the system negatively. However, all subjects stated that the driver has to get some experience to the CMS. The drivers got familiarized with the CMS, so that the small monitor size, the position of the monitors and the changed light conditions for example were perceived as less disturbing. However, several subjects stated that they perceived some risks in terms of road safety during using the CMS.

Objects were perceived smaller on the display. This issue was criticised by all subjects. It is of great importance, especially for manoeuvring, that the displayed image supports the driver in estimating the real object size.

Contrast and colour intensity change depending on sunlight. The contrast between the trailer and road is hardly perceivable. Shadows on the display appear very dark during strong sun exposure, so that objects in the shade of the trailer are not clearly visible. Thus, subjects are not able to estimate the distance to the kerb accurately.

Images on the monitors differ in terms of their colour in case of sun exposure. This resulted in increase glance durations towards the monitor in order to recognize objects on the display. Bright vehicles as well as outlines of distant objects (e.g. head lights of other vehicles) are only poorly or not at all recognisable.

Driving in a roundabout was assessed as "rather disturbing". In the subjects' perception the image jiggled on the monitor and the contrast ratio was so low that it was only possible to estimate the distance between the wheels and the traffic island by close observation.

Nine of ten subjects mentioned that the present weaknesses of the CMS have to be eliminated before the system can be used by customers. With regard to spatial depth perception, the majority of the subjects indicated that they perceived the spatial depth as limited due to the reduction in size of the objects, but the measurement of speed distances would be still possible.

The less positive assessment of the image quality by the subjects in the present study is probably the result of incident solar radiation during the performance of the tests which yielded different levels of glare effects on both monitors. Furthermore, due to the strong shadow formation on the monitors the recognisability of (distant) traffic objects was strongly restricted. Travelling through urban scenarios as well as the estimation of different speeds on motorways was assessed as positive. Besides that, the distance to the pylons was also evaluated as more detailed in CMS than in exterior mirrors.

In addition to technical investigations (avoidance of glare and shadows), further tests should be performed in order to allow for a more comprehensive assessment of the suitability of CMS in trucks. Tests should include a broad range of weather conditions (rain, fog, snow and at night) and investigate the effect of familiarization with the CMS. This test would help to confirm the present statements of the subjects, e.g. with regard to the size and position of the monitors.

6 Conclusions and Recommendations

In general, it was shown that CMS, which meet specific quality criteria, are able to adequately display the indirect rear view to the driver, both for cars and trucks. Depending on the design, it is even possible to receive more information about the rear space from a CMS than is possible with mirror systems. Nevertheless, both solutions show fundamental differences. For example, depth information or a spatial impression of the image is always present in mirrors, however with the CMS this is not possible due to the two-dimensional representation. Furthermore, the field of vision in mirrors can be changed slightly through head movements, this is not possible with the CMS.

In general, the CMS is more resistant to soil and rain drops than the mirror, because the camera is small and the display is installed in the interior. The small camera size is also an advantage with regard to aerodynamics. However, frost, cold and electromagnetic interferences can lead to problems with the CMS. The CMS does not function without power; whereas a mirror is always ready for use. In direct sun exposure the CMS is superior, as it avoids the driver being exposed to direct physical glare. Furthermore, unlike mirrors, it offers the possibility of enhancing or weakening contrasts—depending on the ambient brightness—resulting in an increased comfort of viewing and information content of the image—especially at night. However, the adaptability of the CMS (required time to adapt to differences in brightness and ability to display a large range of brightness levels) is of importance here. Depending on the location of the monitor of the CMS, reflections or glaring can occur on the display. Covers or housings installed at the monitors could remedy this disadvantages. Furthermore, the possibility of artifacts such as "blooming" or "smear" is typical in CMS. This may result in images which do not clearly depict the real conditions, especially when artificial light sources are displayed. On the overall, there is no clear preference for CMS or exterior mirrors, because both systems show advantages and disadvantages. However, the CMS has to meet certain requirements in order to ensure equivalence with the mirrors:

- The electromagnetic compatibility must be ensured
- Good colour and contrast reproduction, minimisation of artefacts
- Quick adaption to changes in ambient brightness
- Representation with no time delay
- Detection and immediate display of image losses or, even better, ensuring that image losses do not occur
- Frost and condensation protection

The test drives with subjects showed that the change from mirrors to CMS does not necessarily result in safety-critical situations, but CMS require a certain period of familiarization. However, the time needed for familiarization seems to be rather short. For cars, with regard to speed and distance estimations, it was found that these are carried out more conservatively with the CMS than with mirrors, i.e. subjects waited for slightly larger gaps before pulling out. As for the trucks, images displayed on the wide-angle mirrors were distorted and represented in relatively small size due to the concavity of the mirror. In comparison to the mirrors the CMS displayed the image more clearly. The reverse driving task was performed more easily with the CMS than when using exterior mirrors. However, distant objects were more difficult to estimate using the CMS due to the lack of depth information. The exact location of the rear part of the vehicle is particularly important when manoeuvring. Here, an additional close-up view of the rear part of the vehicle would be desirable for both mirrors and CMS.

With regard to the positions of the monitors, some subjects stated that information about the left side should always be displayed on the left-hand side. The same applies to the right side. They stated that information does not necessarily have to be displayed close to the A-pillar, but can also be displayed closer to the steering wheel. This was considered positive for truck drivers, because the number of head movements of the driver would be reduced and the details of the images displayed on the monitor would be better recognizable. The area available for direct view increases, if the mirror can be omitted. This was seen as another benefit of the CMS. An installation of the monitors at a location far below the central field of view was considered as undesirable. A sufficiently large representation of the objects must be ensured for recognising distant objects on the monitor. The subjects also found it important that the resolution on the monitor is sufficiently high and comes close to the resolution of the mirrors. A high quality of the colours (especially the white) is also desired. For trucks it was noted that the display should be as large as possible. Far-sighted people should wear glasses when using the CMS, because the monitor of the CMS is positioned closer to the driver than the mirror.

A subject survey showed a medium level of acceptance of the CMS, which did not change when the CMS was used for a longer period of time. It can therefore be assumed that the average expectations of the road users on the CMS were met during the test use.

References

1. Alliance of Automobile Manufacturers (AAM): Statement of principles on human machine interfaces (HMI) for in-vehicle information and communication systems. Version 3.0. Alliance of Automobile Manufacturers, Washington. http://www.umich.edu/~driving/guidelines/AAM_Statement2003.pdf (2003)
2. Bach, P., Rüter, G., Carstengerdes, N., Wender, K.F., Otte, D.: (2006)Untersuchung von Verkehrssicherheitsaspekten durch die Verwendung asphärischer Außenspiegel. [Examination of traffic safety aspects using aspherical exterior mirrors] BASt-Reihe "Fahrzeugtechnik", Heft BASt-F56, Wirtschaftsverlag NW, Bremerhaven. [BASt series "Vehicle Safety", issue BASt-F56, Wirtschaftsverlag NW, Bremerhaven.]
3. Bayerl, S.: Kontextbasierte Vorhersage von kritischen Fahrstreifenwechseln, Masterarbeit, Fachhochschule Deggendorf. (Context-based forecast of critical changing of lanes Master Thesis, University of Applied Sciences Deggendorf) (2012)
4. Bowles, T.S.: Motorway overtaking with four types of exterior rear view mirror. In: International Symposium on Man-Machine Systems, 2, Transport systems and vehicle control. Institute of Electrical and Electronic Engineers (1969)
5. Burger, W.J., Mulholland, M.U., Smith, R.L., Sharkey, T.J.: Passenger vehicle, light truck an van convex mirror optimization and evaluation studies 1 Convex mirror optimization. Department of Transportation, National Highway Traffic Safety Administration, Washington (1980)
6. Denson, S. (2013): Kamera-Monitor-Systeme als Ersatz für Außenspiegel–Experimentelle Untersuchung von Wahrnehmungsaspekten in der Fahrer-Fahrzeug-Interaktion. [Camera-monitor systems as replacement for exterior mirrors–Experimental investigation of perception aspects in driver-vehicle interaction.] Master thesis. Institut für kognitive Neurowissenschaft, Abteilung für Neuropsychologie, Ruhr-Universität Bochum. [Institute for Cognitive Neuroscience, Department of Neuropsychology, Ruhr University, Bochum]
7. de Vos, A.P.: Non-planar Driver's Side Rear-View Mirrors: A Survey of Mirror Types and European Driver Experience and a Driver Behaviour Study on the Influence of Experience and Driver Age on Gap Acceptance and Vehicle Detection. Final Report. TNO Human Factors, Soesterberg (2000)
8. Dube, C.: Bewegungswahrnehmung und Wahrnehmung bewegter Bilder (Motion Pictures). [Perception of Motion and Perception of Moving Images (Motion Pictures)] GRIN Verlag (2011)
9. Fisher, J.A., Galer, I.A.R.: The effects of decreasing the radius of curvature of convex external rear view mirrors upon drivers' judgements of vehicles approaching in the rearward visual field. Ergonomics **27**, 1209–1224 (1984)
10. Flannagan, M.J.: Current status and future prospects for non-planar rear-view mirrors. In: SAE Publication. Human Factors (2000)
11. Flannagan, M.J.: Distance Perception in Camera-Based RearVision Systems. Michigan, USA (2002)
12. Flannagan, M.J., Mefford, M.L.: Distance perception with a camera-based rear vision system in actual driving. In: Driving Assessment 3rd International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design (2005)
13. Flannagan, M.J., Sivak, M., Simpson, J.K.: The relative importance of pictorial and nonpictorial distance cues for driver vision. In: Proceedings of the International Driving

Symposium on Human Factors in Driving Assessment, Training and Vehicle Design, Aspen, CO, pp. 214–218 (2001)
14. Flannagan, M.J., Sivak, M., Traube, E.C.: Driver perceptual adaptation to nonplanar rear-view mirrors. In: Serafin, C., Zobel, G.(Hrg.) (eds.) Automotive Design Advancements in Human Factors: Improving Drivers' Comfort and Performance, pp. 213–220. Society of Automotive Engineers, Warrendale (1996)
15. Flannagan, M.J., Sivak, M.: Distance cues and Fields of View in Rear Vision Systems. SAEPuplication Detroit, Michigan (2006)
16. Goldstein, E.B.: *Wahrnehmungspsychologie. [Psychology of Perception]* Manfred Ritter (Publisher) Heidelberg: Spektrum Akademischer Verlag, (2nd edition) (2002) (Chapter 1–2 & 5–9)
17. Henderson, L., Smith, R.L., Burger, W.J., Stern, S.D.: Visibility from motor vehicles. SAE Technical Paper. Pennsylvania: Society of Automotive Engineers, Warrendale (1983)
18. Henning, M.J., Georgeon, O., Wynn, T., Krems, J.F.: Modeling driver behaviour in order to infer the intention to change lanes. In: Brusque, C. (ed.) Proceedings of the European Conference on Human Interface Design for Intelligent Transport Systems, pp. 113–120. Humanist Publications, Lyon (2008)
19. Henning, M.: Das dynamische Arbeitsgedächtnis beim Spurwechsel. [Dynamic working memory when changing lanes] Presentation at the German Aerospace Centre, Institute of Transportation Systems, Braunschweig (2009)
20. Henning, M.J.: Preparation for lane change manoeuvres: Behavioural indicators and underlying cognitive processes. Dissertation, Technical University Chemnitz (2010)
21. Hoffmann, H., Wipking, C., Blanke, L., Falkenstein, M.: Experimentelle Untersuchung zur Unterstützung der Entwicklung von Fahrerassistenzsystemen für ältere Kraftfahrer. [Experimental Study for the Support of the Development of Driver-Assistance Systems for older drivers] Reports by the Federal Highway Research Institute, Vehicle Technology, Issue F 86 (2013)
22. ISO 16505 (Road vehicles—Ergonomic and performance aspects of Camera Monitor Systems – Requirements and test procedures (2015)
23. Liebner, M., Klanner, F., Stiller, C.: Der Fahrer im Mittelpunkt–Eye-Tracking als Schlüssel zum mitdenkenden Fahrzeug In: [Drivers at the Centre of Eye-Tracking as a Key for Intelligent Vehicles, in:] 8. Workshop Fahrerassistenzsysteme (FAS2012). [Driver-assistance system workshop (FAS2012)] Walting, Germany, pp. 87–96 (2010)
24. Marandi, D.: Analyse und Bewertung eines Kamera-Monitor-Systems als Ersatz für Außenspiegel an Kraftfahrzeugen". [Analysis and Evaluation of a Camera-Monitor System as Replacement for Exterior Mirrors in Vehicles] Bachelor thesis. Department of Mechanical Engineering and Process Engineering, Niederrhein University, Viersen (2013)
25. Mortimer, R.G.: The effects of convex exterior mirrors on lane-changing and passing performance of drivers. SAE Technical Paper Series. Society of Automotive Engineers, New York (1971)
26. Mortimer, R.G., Jørgensen, C.M.: Drivers' vision and performance with convex exterior rear-view mirrors. SAE Technical Paper. Society of Automotive Engineers, New York (1974)
27. Rauch, N.: Ein verhaltensbasiertes Messmodell zur Erfassung von Situationsbewusstsein im Fahrkontext. [A Behaviour-Based Measurement Model for the Capturing of Situational Awareness in Driving]. Dissertation. University of Würzburg (2009)
28. Rickheit, G., Herrmann, T., Deutsch, W.: Psycholinguistik/ Psycholinguistics: Ein internationales Handbuch, [An International Manual] Walter de Gruyter. 24 (2003)
29. Robert Bosch GmbH (Publisher): Kraftfahrtechnisches Taschenbuch [Auto-technical Pocket Book] Wiesbaden: Friedr. Vieweg & Sohn Verlag (2007)
30. Suguira, D., Kimura, K. Outside rear-view mirror requirements for passenger cars—curvature, size, and location. SAE Technical Paper. Society of Automotive Engineers, Warrendale (1978)
31. TÜV Report: Untersuchung einer Kamera-Monitor-Einrichtung als Ersatz für die Außenrückspiegel Gruppe III nach ECE-R 46. [Examination of a Camera-Monitor Device

as Replacement for Group III Exterior Mirrors according to ECE-R46]. Report No.: 21183041-001 (2013)
32. UN-R-46: Regulation No 46 (R 46) of the United Nations Economic Commission for Europe (UN/ECE)—Uniform provisions concerning the approval of devices for indirect vision and of motor vehicles with regard to the installation of these devices. http://www.bmvbs.de/cae/servlet/contentblob/44908/publicationFile/41741/r-46-rueckspiegel-pdf.pdf (2010). Accessed: 20 July 2013 (14:41)
33. Wagner, P. (2014), Internet research. http://www.filmscanner.info/Glossar_B.html#Blooming (2014). Accessed 1 Dec 2014
34. Walraven, P.L., Michon, J.A.: The influence of some side mirror parameters on the decisions of drivers. SAE Technical Paper. Society of Automotive Engineers, New York (1969)
35. WIKI: http://de.wikipedia.org/wiki/Smear (2014a). Accessed 1 Dec 2014
36. WIKI: http://de.wikipedia.org/wiki/Lens_Flare (2014b). Accessed 1 Dec 2014
37. Zwahlen, H.T., Adams, C.C., Debald, D.P.: Safety Aspects of Crt Touch Panel Controls in Automobiles, pp. 335–344. Elsevier, Amsterdam (1988)

CMS Concept for Commercial Vehicles: Optimized Fuel Efficiency and Increased Safe Mobility

Tobias Schmalriede

Abstract This article discusses a mirror replacement system designed for commercial vehicles. According to Continental AG (CAG), four megatrends are emerging in the automotive industry. Every successful automotive product must serve at least one of these trends. In the case of camera monitor systems (CMS), the demands of two trends are met: safe and clean mobility. Based on several aerodynamic studies a side-view mirror replacement can reduce the fuel consumption of up to 2.9 %. Besides the field of view (FOV) restrictions current mirror technologies have to observe the surroundings, the article also outlines forward-FOV obstructions. Limitations of conventional mirrors under low light conditions are also mentioned as well as challenges camera monitor systems have. In this article the author discusses technical limitations as well as application-specific demands by illustrating the advantages for drivers of camera monitor systems over conventional mirrors. The technical concept of the ProViu®Mirror outlined in the article is CAG's version of a camera monitor system.

Keywords CMS concept · ProViu®mirror · Commercial vehicle · Aerodynamic improvements · Camera monitor system · Side-view mirror replacement · Mirror replacement

List of Abbreviations

ADAS	Advanced driver assistance systems
ASIL	Automotive safety integrity level
CAG	Continental AG
CMS	Camera monitor system
FOV	Field of view
ISO	International Organization for Standardization
IP	International Protection
MTF	Modulation transfer function

T. Schmalriede (✉)
Continental-Corporation AG, Frankfurt am Main, Germany
e-mail: tobias.schmalriede@continental-corporation.com

A. Terzis (ed.), *Handbook of Camera Monitor Systems*,
Augmented Vision and Reality 5, DOI 10.1007/978-3-319-29611-1_13

MTF10P	Modulation transfer function 10 percent
NRE	Non-recurring expenses
TFT	Thin-film transistor
μC	Microcontroller

1 Automotive Megatrends

CAG has identified four megatrends (see Fig. 1) currently faced by the automotive industry. These trends will influence every successful new product.

- Safe mobility
- Intelligent driving
- Clean power
- Global mobility

The side-view mirror replacement ProViu®Mirror serves two of these megatrends:

- Safe mobility
- Clean mobility

1.1 Safe Mobility

Safely maneuvering vehicles requires a high level of awareness of the obstacles surrounding the vehicle. Indirect vision devices reduce blind spots and thus increase

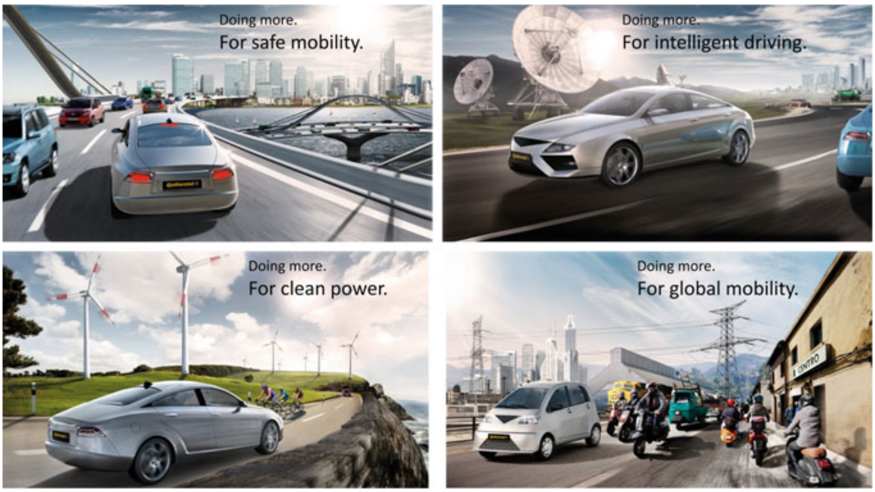


Fig. 1 Continental's four megatrends [15]

the amount of relevant data available to the driver. Unfortunately, physical limitations mean that not every blind spot can be covered. Camera monitor systems (CMS) are a new technology with the potential to further minimize blind spots.

In 2012, a representative year, 3,402 accidents between commercial vehicles (trucks and agricultural vehicles) and cyclists occurred in Germany. Furthermore, 55 collisions between pedestrians and turning commercial vehicles occurred in German inner cities. It may reasonably be assumed that these accidents are caused by insufficient information for the driver (blind spots) [1]. An increased field of view (FOV) reduces fatal accidents, especially with cyclists or pedestrians [1] and [2].

Forward-directed FOV is limited by increasingly large mirror systems [3]. By placing the display of a CMS in front of the A-pillar, blind spots can be further reduced (see Fig. 2).

Even though the relevant FOV can be covered by an indirect vision device, challenging environmental conditions may decrease the information quality. This is especially problematic when driving at night or when the sun is directly reflected in the mirror. In such circumstances the driver is blinded and distracted while driving the vehicle.

It has been shown in field studies that enhanced image quality can increase driving comfort, particularly in conditions with poor lighting. It is easier to identify obstacles on a well-lit image [4].

In addition to the practical advantages of replacing side-view mirrors with camera monitor systems, such systems also have strategic importance for the future. Sensors on the side of the vehicle will convey essential information for autonomous driving in the future. Functions such as blind spot detection, trailer support, lane detection, lane change assist and turn assist are steps on the path to autonomous driving.

Fig. 2 Safety benefits of a camera monitor system

1.2 Clean Mobility

As the world is running out of fossil fuels [5], the demand for enhanced efficiency is constantly increasing. Furthermore, global statutory gCO_2/km requirements are becoming stricter [6] in order to protect the environment and increase quality of life, especially in larger cities. Lower fuel consumption leads to commercial mobility that is more environmentally friendly.

ProViu®Mirror, a CMS for replacing side-view mirrors, improves vehicle aerodynamics, which in turn reduces fuel consumption by about 2 % due to less aerodynamic resistance. This has the potential to reduce fuel costs by about €1,300 annually [3].

In combination with other aerodynamic improvements, replacing side-view mirrors can reduce fuel consumption by up to 2.9 %. "While the mirrors on the test truck were of limited relevance to aerodynamics when located in negative pressure areas on the sides, after the changes to the front end they were directly exposed to the air flow and increased drag by as much as 10 %. The use of cameras instead of traditional mirrors could reduce fuel consumption by an additional 2.9 %" [7].

As mirrors rely on physical limitations (angle of incidence equals angle of reflection), they need to be directly visible (i.e. through the windows). Unfortunately, these areas are characterized by high air flow velocities. Vehicle components located in these areas create significant noise [8].

Cameras, on the other hand, can overcome this limitation by being placed in other locations (e.g. hiding the camera modules in the cabin roof design). Even though some FOV can only be covered if the camera is placed at the same location as the traditional mirrors, noise creation will be reduced due to the smaller size of the cameras [9].

Another advantage of CMS is that they reduce dirt buildup on windows. Not only can the indirect vision device (cameras instead of mirrors) be cleaned less frequently, the windows also need less cleaning due to the fact that it is not required anymore to perceive tiny details in the outside located mirrors. This leads to lower costs of ownership and an improved truck appearance [10].

2 Challenges

Besides the significant benefits camera monitor systems have to offer vehicle drivers and owners, certain limitations need to be considered. ProViu®Mirror targets the following challenges.

2.1 Providing the Right Information at the Right Time

Currently trucks are required to have six different mirrors to cover all legally required FOV [11] according to UN ECE R46. These legally mandated FOV could be increased in certain driving situation to reduce blind spots. In some driving situations the driver needs additional information. Several studies identified additional FOV needed in certain driving situations to safely maneuver commercial vehicles.

According to a head movement survey done in cooperation with Volvo, the legally required class II mirrors are extended by 24° on the passenger side in certain driving conditions by using head movements. For class IV mirrors it is rare that no head movement is used [10].

A field study conducted in cooperation with MAN identified the need to combine different FOV into one display to have the entire scene available at a single glance. Using multiple mirrors is a strain on the driver due to the number of different mirrors around the cabin and the challenge of looking into the right mirror at the right time [12].

Studies on drivers' visibility situations came to the conclusion that for brief periods of time drivers need a different FOV and that the FOV should be panned rather than enlarged. A study in cooperation with Daimler Trucks proposed a FOV functionality based on vehicle parameters such as velocity and steering wheel angle [3].

2.2 Technology

One of drivers' important tasks is estimating the distance of oncoming vehicles. This is relevant when passing other vehicles, for example. While the driver has better depth perception when looking in a traditional mirror, a monitor does not deliver this information [10]. The driver relies on information such as size, size increase, movement of vehicle geometries or light sources [4]. Accordingly, it is very important to properly reproduce these conditions.

The ProViu®Mirror camera monitor system uses high dynamic range imagers to reproduce point light sources as accurately as possible (see Fig. 3).

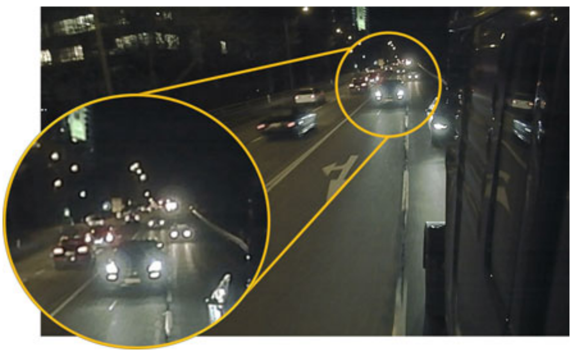


Fig. 3 Point light sources

Regarding displays, automotive-qualified displays are often designed for applications in landscape format, such as instrument clusters or center stack displays in the S-Class [13]. Turning these displays by 90° (as is the case for side-view mirror replacements) will lead to significant reductions in brightness when using polarized sunglasses (see Fig. 4); the display emits linearly polarized light which is almost completely cancelled out as soon as it reaches linearly polarized sunglasses [14].

A special thin-film transistor (TFT) display with rotated polarization is used for the ProViu®Mirror. The polarization is oriented in such a way that the use of polarized sunglasses is possible (see Fig. 5).

Fig. 4 Use of improper display polarization

Fig. 5 Polarization orientation aligned

3 Solution

ProViu®Mirror is a platform camera monitor system which can be adapted to meet customer needs. In general, it consists of two 12.3″ displays and four camera modules per vehicle. Two camera modules on each side are built into design-optimized housing and mounted onto the cabin's roof (see Fig. 6). The 12.3″ displays are mounted in a portrait orientation in front of the A-pillar (see Fig. 7). Since 2012, ProViu®Mirror has been constantly improved based on field test experience.

3.1 System Architecture

The system architecture (see Fig. 8) is designed to include a display and two cameras on each side. They are connected via an LVDS connection, which enables low latency. The split screen and camera initialization is realized in the display, meaning that no additional ECU is required. The bi-directional control channel between the system ECU and camera modules to command the imagers is implemented by LVDS back channel. Camera firmware can also be updated via the LVDS backchannel.

Fig. 6 CMS cabin design proposal

Fig. 7 Cabin interior design proposal

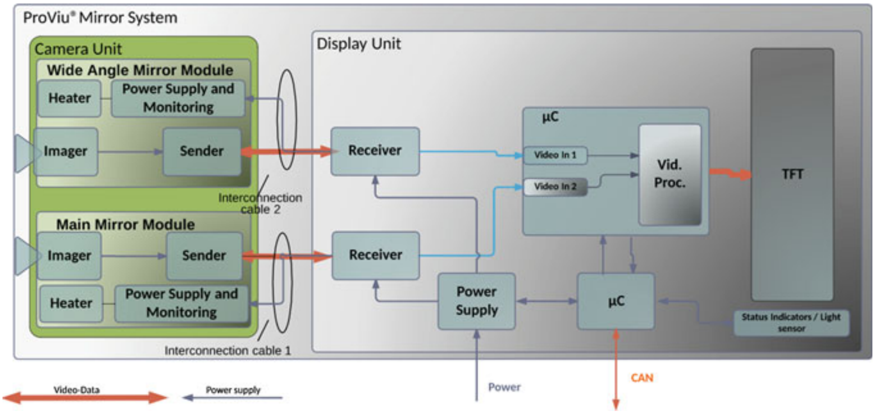


Fig. 8 ProViu®Mirror system architecture

With the experience of more than 10 recent ASIL-relevant projects, the architecture is designed to comply with ASIL requirements. Furthermore, the entire development project is realized in an ASPICE level 3 project. This approach is necessary to develop a proper camera monitor system that is relevant to safety.

One of the most outstanding advantages of this concept is that neither a CAN firewall nor a CAN traffic limiter is required because no µC is used outside in the camera which can be hacked.

3.2 *Display*

The display used is a 1920 px × 720 px high resolution 12.3″ display. It is designed to be mounted on different A-pillar forms (see Fig. 9). This allows the OEM to

Fig. 9 Proposed installation in the A-pillar

design cabins with improved forward direct sight. The display has an anti-glare, anti-reflection coating to preserve readability even in the most intense sunlight.

The required FOV are displayed in a split screen arrangement, similar to traditional mirrors and as expected by drivers. The TFT panel is a mature display component which is also used in infotainment and instrument cluster applications. As one of the largest automotive suppliers in the world, CAG has access to a variety of TFT panels and can adapt the TFT to customers' needs [15].

The depth of the top of the display is reduced for straightforward integration into the A-pillar.

3.3 Camera

IP69k miniature camera modules with pigtails provide designers with the freedom to attach the cameras anywhere in the cabin. If a camera fixture is requested, ProViu®Mirror offers a foldable camera arm to place the camera modules. The camera design housings are vehicle specific and optimized to improve the aerodynamic of the entire vehicle (Fig. 10).

The high resolution and high dynamic range imagers have been improved to capture the surroundings of the vehicle. These cameras deliver a high-quality image of the environment. This technology has been successfully implemented in several applications, such as 360° surround view. The in-house camera module production is working closely with our imager and optics experts to provide drivers with an optimal driving experience. The lens opening angle can be adapted to customer

Fig. 10 Proposed installation in the roof

needs. Typically 60° and 100° horizontal FOV are provided. CAG works in close cooperation with the OEM to determine the optimal trade-offs between maximum opening angle and the desired magnification factor in order to eliminate blind spots and maximize object sizes on the display.

The aerodynamic and design-optimized housing can be adapted to different cabin roofs to seamlessly fit into the appearance of the cabin design.

It is recommended that the camera modules be installed in high positions to reduce soiling and far to the front to reduce the required lens opening angle.

3.4 Features

ProViu®Mirror offers a variety of features to support the driver while maneuvering the vehicle. Its basic function is the image stream presentation in a split screen arrangement.

With the development tool chain, a group of required FOV (see Fig. 11) can be set-up and tested.

Side-view mirror replacement solutions are relevant to ASIL [16]. No general ASIL level is defined in ISO16505:2015. In collaboration with the OEM, CAG will adjust the system concept according to the use of the system inside the specific vehicle model. Critical points requiring further specific adaptation to the OEM architecture

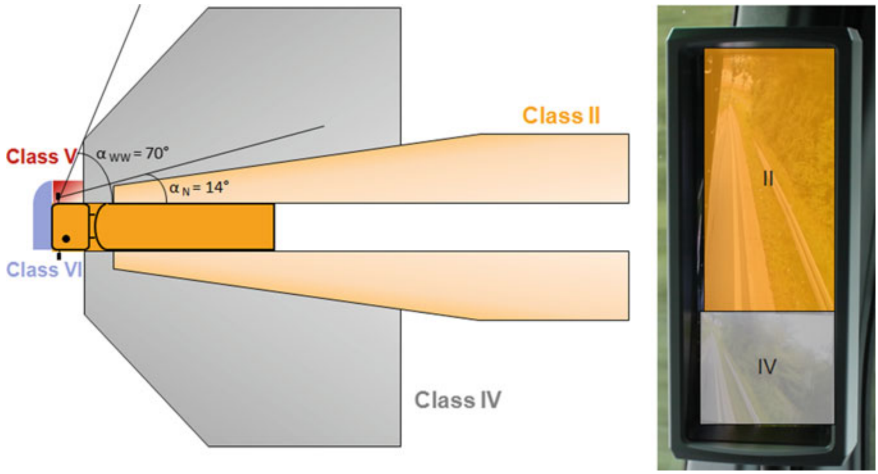


Fig. 11 Field of view coverage and display arrangement

can then be identified. The updated system concept will address the safety goals which result from this evaluation. The currently available ProViu®Mirror basic system fulfills legal requirements including latency and MTF10P.

Upcoming features will certainly include:

- Turn assist
- Trailer hitch assist
- Blind spot warning
- Lane recognition
- Lane change assist

3.5 Variations

In comparison with the passenger vehicle market, most truck vehicle models have smaller volumes which increase the necessity for a platform approach with low non-recurring expenses (NRE) for research and development as well as manufacturing equipment [17].

ProViu®Mirror offers truck- and bus-specific adaptations; different vehicle widths are supported by a scalable camera housing approach. Additional camera support can be provided for bus applications.

4 Conclusion

ProViu®Mirror enables the OEM to optimize the vehicle geometries to reduce the fuel consumption by up to 2.9 %. It enhances the safety by increasing the image quality under low light conditions as well as reducing disturbing intense head beam lights from vehicles behind.

Challenges drivers currently facing, such as obtaining the right information at the right time can be overcome by camera monitor systems. ProViu®Mirror is capable to show a larger FOV than traditional mirrors.

Constantly improving technologies like the increase in resolution of imagers and display as well as imager functionalities such as high dynamic range increased the driver acceptance over the last years significantly.

The proposed system architecture of ProViu®Mirror allows the OEM a flexible mechanically as well as electronically integration into the vehicle.

References

1. Pöppel-Decker, M., Schreck, B.: Unfallgeschehen zwischen rechtsabbiegenden Güterkraftfahrzeugen und geradeausfahrenden Radfahrern,. Zeitschrift für Verkehrssicherheit, pp. 239–242 (2014)
2. Platzer, G.: The geometry of automotive rearview mirrors - why blind zones exist and strategies to overcome them. SAE Technical Paper Series, Vol. 950601 (1995). (1 Februrary)
3. Bothe, A.: Analyse dynamischer Sichtsituationen zur ergonomischen Auslegung von Kamera-Monitor-Systemen (KMS) in schweren Nutzfahrzeugen. Fachbereich Maschinenbau der Technischen Universität Darmstadt, Darmstadt (2014)
4. Betke, M., Nguyen, H.: Highway scene analysis from a moving vehicle under reduced visibility conditions. In: IEEE International Conference on Intelligent Vehicles, stuttgart (1998)
5. Shafiee, S., Topal, E.: When will fossil fuel reserves be diminished? *Energy Policy*, pp. 181–189 (2009 Jan 08)
6. An, F., Sauer, A.: Comparison of Passenger Vehicle Fuel Economy and GHG Emmision Standards Around the World. Pew Center on Global Climate Change, Arlington (2004)
7. Devesa, A., Indinger, T.: Verbrauchsreduktion an Nutzfahrzeugkombinationen durch aerodynamische Maßnahmen. Forschungsvereinigung Automobiltechnik e.V, Berlin (2011)
8. Helfer, M.: Aeroakustische Messungen an Kraftfahrzeugen in Windkanälen. Stuttgart (2004)
9. Tenstam, A.: CFD simulations of aeroacoustic generation of noise from A-post/wing mirror. Volvo GTT, Gothenburg (2008)
10. Fagerström, K.F., Gardlund, A.: Mirror Replacement in Trucks. Chalmers University of Technology, Göteborg (2012)
11. UNECE. *Uniform provisions concerning the approval of devices for indirect vision and of motor vehicles with regard to the installation of these devices* (Addendum 45 Revision 5 ed.). Geneva, (2013)
12. Zaindl, A.: Spiegelersatz am Nutzfahrzeug. Ergonomie Aktuell **16**, 30–33 (2015)
13. Daimler A.G.: *S-Klasse Preisliste*. Retrieved 09 01, 2015, from http://www.mercedes-benz.de/content/media_library/germany/mpc_germany/de/mercedes-benz_deutschland/pkw_emb_nextgen/neufahrzeuge/s-klasse/preislisten_pdf/0.object-Single-MEDIA.download.tmp/Preisliste_S-Klasse_150407_02.pdf (2015)

14. Haferkorn, H.: Optik: Physikalisch-technische Grundlagen und Anwendungen, 4th edn. Wiley-VCH Verlag GmbH & Co, KGaA (2002)
15. Aktiengesellschaft, Continental: Fact Book 2014. Continental Aktiengesellschaft, Hannover (2015)
16. ISO. *ISO16505*—Road vehicles—ergonomic and performance aspects of camera-monitor systems—requirements and test procedures. Geneva (2014)
17. OICA. World Motor Vehicle Production. Retrieved 08 25, 2015, from http://www.oica.net/wp-content/uploads//ranking-2013s-2.pdf(2013)

Part V
Advanced Topics

Optimization of Demanding Scenarios in CMS and Image Quality Criteria

Mark Leznik and Anestis Terzis

Abstract The number of camera-based advanced driver assistance systems is growing steadily. Starting with rear-view cameras and ranging to complete 360° surround view systems. Such systems are not limited to displaying the captured video signals but also permit additional functions such as object detection based on the video signals. In the future, such systems will become mandatory, as have the rear-view cameras for new passenger cars in the US market. Another example is the new international standard ISO 16505 that describes the replacement of mandatory vehicle mirrors with camera monitor systems. These new technologies can be the base of future innovations for indirect vision. In order to evaluate and compare such systems, specific emphasis has to be given to the appropriate Image Quality (IQ) criteria in addition to the normative criteria.This work introduces a prototype vehicle setup utilizing a High Dynamic Range (HDR) camera for a rear-view driver assistance application. While the system shows promising results in daylight conditions, there are still several issues that are of concern when it comes to twilight, low-light, and night applications. A newly defined IQ assessment set is used for a real-time analysis of laboratory and outdoor scenes. Further, an optimization using the implemented modules in an automotive framework is performed and presented.

Keywords High dynamic range · Image quality for CMS · Entropy and noise · Image quality assessment · Automotive imager

List of Abbreviations

AACDG	Average Absolute Color Deviation from Greyscale
ADTF	Automotive Data and Time-Triggered Framework

M. Leznik (✉)
Ulm University of Applied Sciences, Ulm, Germany
e-mail: mark.leznik@gmail.com

A. Terzis
Department of Electrical Engineering and Information Technology,
Ulm University of Applied Sciences, Ulm, Germany
e-mail: terzis@hs-ulm.de

A. Terzis (ed.), *Handbook of Camera Monitor Systems*,
Augmented Vision and Reality 5, DOI 10.1007/978-3-319-29611-1_14

AWB	Auto White Balancing
CCD	Charge-coupled device
CMOS	Complementary Metal Oxide Semiconductor
CT	Color Temperature
FPGA	Field Programmable Gate Array
HDR	High Dynamic Range
HVS	Human Visual System
IQ	Image Quality
LDR	Low Dynamic Range
LED	Light-Emitting Diode
LVDS	Low Voltage Differential Signaling
MSE	Mean Squared Error
MTF	Modulation Transfer Function
PSNR	Peak Signal-to-Noise Ratio
RAM	Random Access Memory
RGB	Red Green Blue

1 Introduction

The most prominent issues in automotive imaging technology, especially at night, are high image noise and color genuineness, as well as haloing effects from car headlights. Decreasing the framerate, and by that increasing the available exposure time for the imager is not always an option in a safety relevant automotive application. Noise, especially in homogeneous regions is very present as a consequence. Color genuineness is of high importance as a perception criteria for the driver, e.g. the distinguish-ability and differentiation of head and tail lights and also traffic lights.

While the utilized HDR hardware already provides a better image as a conventional Low Dynamic Range (LDR) camera, it is not exempt from IQ issues and, in fact even develops some new of its own. The application of this hardware in automotive conditions also presents new issues to be tackled.

However, before any optimization and improvement of the IQ with consideration of the described issues can be performed, a clear definition of criteria using which it is measured must take place.

Using the defined criteria, IQ optimizations are performed, measured, evaluated and of course, presented.

This work is organized as follows: Sect. 2 provides the background information on HDR image acquisition and display. Section 3 describes the applied HDR camera. It also introduces the Automotive Data and Time-Triggered Framework (ADTF) which was used to develop the real-time IQ assessment criteria described in Sect. 5. Section 4 gives some insights into the IQ issues that had to be dealt with in this work, and provides explanations as to why those issues occur. Several approaches to optimize the performance of the image sensor and evaluate camera behavior before and after the improvement process are shown in Sect. 6. Section 7 concludes this work with a summary and some ideas for possible future developments.

2 HDR Imaging and Display

Dynamic range as it is should be discussed before any extensive use of the term. Reinhard [1] proposes several definitions for the dynamic range. A dynamic range of an image would be the ratio between the lightest and darkest pixels. For a displaying device, the dynamic range would be the ratio between the maximum and minimum of the emittable luminance. The dynamic range of a camera is defined as the luminance that just saturates the sensor and the luminance that lifts the camera response to one standard deviation above noise floor [2]. The dynamic light range on earth can be calculated using the average light of the sun and the average light of the stars, and sums up to approximately 160 dB. While a conventional, or LDR image is able to capture about 60 dB [3], a high dynamic range image, captured either using several exposures of a low dynamic range camera or a dedicated high dynamic range camera, is able to capture 120 dB (see Fig. 1 for a comparison). It is comprehensible why high dynamic range imaging is the clear choice for future automotive camera generations. A camera as it is can be regarded as an imperfect device for measuring the radiance distribution of a scene, in that it cannot capture the full spectral content and dynamic range [1]. But why is that? A typical camera lens is used to focus the incoming light onto the image sensor. Technically, the information is available at this point; however it is not being captured due to the limitations of an image sensor. As Chaurasiya et al. [4] remark, the problem of HDR imaging is two-fold: firstly, the capture of the scene's luminance must be mastered, and secondly, the reproduction of the captured information as genuine as possible on an LDR display device. An HDR image is generally taken by capturing multiple, differently exposed photographs of a scene and constructing the HDR radiance map using the camera's response function. Assuming the camera's response functions is known, this is done by dividing each pixel by its exposure time [4]. Using this approach, it is possible to capture the dynamic range of an HDR image, which covers a wider luminance spectrum than a conventional LDR image. Since conventional displays are unable to display such HDR images, a compression with preserved visibility is being performed. This term is better known as tone-mapping or perception-based tone reproduction. The following sections shall describe the image acquisition and later on, display in detail.

Fig. 1 From *left* to *right*, the first two images were captured with a conventional low dynamic range camera, with long and short exposure accordingly, compared to the last image taken with a high dynamic range camera

2.1 HDR Image Acquisition

A digital image is being taken by exposing an image sensor to light for a particular amount of time. Technical limitations do not make it possible to capture all light intensities in the scene in one exposure. By taking multiple differently exposed images of a scene, and merging this information, an HDR radiance map containing all the light intensities can be created using a standard LDR camera. Thereafter, a tone-mapped image using the radiance map can be composed. As already established, several images with different exposures are necessary. Meaning, several images with different amount of light projected onto the camera's image sensor. Since those photographs are captured over some time, the scene has to be perfectly still, otherwise, the images must be aligned. As stated before, a camera can be regarded as a light measurement device, however, an imperfect one. Cameras do not respond to light linearly, which means if the value of any given pixel is twice as high compared to any other pixel, this does not necessarily mean that its luminance is also twice as high. Therefore, the multiple exposures cannot be merged to a radiance map in a trivial way. Thus, the next step of the process would then require the calculation of the camera's light response curve. Using the camera's response function and the multiple differently exposed images, the radiance map of a scene can be constructed [4].

2.1.1 Capturing Multiple Exposures

There is no ideal amount of images required to capture an HDR image, since a fair amount of the captured pixels of the multiple images will be either over- or underexposed, the perfect approach would be to capture an infinite amount of differently exposed images. It should however be noted, that while capturing the images required for an HDR image calculation, the different exposure times shall be chosen in a fashion that each and every pixel will be neither over- nor underexposed at least once.

2.1.2 Image Matching and Alignment

Even when using a tripod and capturing a still scene, some movement in the scene itself may occur. This movement demands an alignment of the different exposures between each other. There are several techniques available for this operation, most of them based on computer vision. Chuarasiya et al. [4] propose using projective geometry transformation introduced by Stan Birchfield [5], which however, can handle only slight movement between the exposures. Reinhard [1] suggests two approaches for this task. A technique by Kang et al. [6] which is able to account for both camera and image movement is based on a motion estimation. Both techniques must be performed offline and are quite resource intensive. To overcome this issue, the camera utilized in this work has a unique design, which makes it possible to capture two different exposures at nearly the same time. Thus, motion artifacts between the two exposures are virtually not present and both exposures can directly be merged, without a computational expensive motion compensation.

2.1.3 Camera Response Curve

While not provided by the manufacturer, the camera's light response function can however be measured and estimated. There are several techniques for this process described below. Debevec and Malik [7] developed a technique following the idea that when regarding several pixel over different exposures, parts of the response curve can be constructed. The more difficult task in this case is to combine this information to a singular graph, which is performed using linear optimization. Mitsunaga and Nayar [8] introduce an approach which is suitable when dealing with lower-cost equipment, where the aperture and shutter speed are unknown. They do so using a polynomial approximation to the response function. While the response curve deduction can be performed using every pixel, this would add longer calculation times, which can be reduced using cleverly chosen pixels. Reinhard [1] proposes either choosing samples from images by a provided algorithm, or using just the image histograms as a measure.

2.1.4 Weight Factoring

Over- and underexposed pixels from the multiple images used to construct an HDR image shall not be used in the process. Assuming once again a perfectly still scene with aligned images and a known camera response function, one would choose images by their exposure; however, the problem of which pixel to choose over the other must be solved. Basically, the approaches introduced for this problem all rely on weighting certain pixels over the other. A simple hat function:

$$w(v) = \begin{cases} v - v_{min}, \, for\, v \leq \frac{1}{2}(v_{min} + v_{max}) \\ v_{max} - v, \, for\, v > \frac{1}{2}(v_{min} + v_{max}) \end{cases} \tag{1}$$

where w is the weighting function, v is the digital value and v_{min}, v_{max} are minimum and maximum values, is used by Devebec and Malik [7]. They assume that mid-range pixels are more reliable.

A weighting function is suggested by Mann and Picard [9] and is based on the idea that a higher response sensitivity would mean a more “trustable” pixel, to do that they use a derivative of the camera’s response function:

$$w(v) = \frac{1}{\frac{d}{dv}(log_{10}(f^{-1}(v)))} \tag{2}$$

where w is the weighting function, v is the digital value and f^{-1} is the inverse response function. Using signal theory as suggested by Mitsunaga and Nayar [8] is also possible.

2.1.5 The HDR Radiance Map

Assuming for simplicity, that a camera’s response function is linear, a radiance map can be constructed using the equation provided by Reinhard [1], excluding over-and underexposed pixels, where weight factors are given by $w(Z_{ij})$ for 8-bit values $Z_{ij} = L_e(i,j) = E_e(i,j)\Delta t_k$ at location (i,j) computed from N exposures:

$$L_{ij} = \sum_{k=1}^{N} \frac{Z_{ij} w(Z_{ij})}{\Delta t_k} / \sum_{k=1}^{N} w(Z_{ij}) \tag{3}$$

The exposure time is given by Δt_k for exposure k.

Assuming a known camera response function f given by $Z_{ij} = f(L_e(i,j)) = f(E_e(i,j)\Delta t_k)$, one would calculate the radiance map by inverting the function to calculate exposures $E_e(i,j)\Delta t_k$ from pixel values Z_{ij} which leads to a new equation for an HDR radiance map:

$$L_{ij} = \sum_{k=1}^{N} \frac{f^{-1}(Z_{ij}) w(Z_{ij})}{\Delta t_k} / \sum_{k=1}^{N} w(Z_{ij}) \tag{4}$$

The described process provides a radiance map which contains all the light intensities of a given scene. This information however, cannot be directly shown on a regular display device. In the following, the mapping process of the radiance map pixel values to a conventional display shall be described.

2.2 HDR Image Display

Using previously described processes, it was shown how to capture and then calculate high dynamic range image radiance maps that contain the broad range of illumination of a given scene. Most display devices however are not capable of reproducing this kind of dynamic range. A process, described by Reinhard [1] as the reverse adaptation model, more commonly known as tone-reproduction or tone-mapping, is necessary. This second part of the HDR imaging problem also involves quantization, namely, the correct scaling of the luminance values to the display device.

2.2.1 Photoreceptor Adaptations

The solution used by the Human Visual System (HVS) in this case is quite easy, an adaptation to the dominant light conditions is performed, as an example, Reinhard [1] proposes the reaction towards car lights: while they may not disturb us during the daytime, at night time they are far more perceptible and as a side-effect: uncomfortable. Basically, the human vision system scales luminance levels in a way that lets us distinguish differences within a scene. The human vision system's photoreceptors consist of two types: rods and cones. Rods are very sensitive to light, and are generally used at lower level light conditions, while cones are less sensitive to light and are used during daytime. The illumination is divided by the rods and cones into photopic and scotopic ranges. In an HDR dynamic range, with both dark and bright areas, rods and cones are active, this range is called the mesopic range [1]. The HVS then operates using so-called just noticeable differences. It determines the stimulus, the intensity difference which the eye can distinguish and then scales the scene appropriately. The adapted response function for this approach is the following:

$$\frac{R}{R_{max}} = \frac{I^n}{I^n + \sigma_b^n} \tag{5}$$

where R is the photoreceptor response, R_{max} is the maximum response, I is the light intensity, σ_b is the background intensity and n the sensitivity control exponent, which is generally set to a value between 0.7 and 1.0 [1].

2.2.2 Perception Based Tone Reproduction

A number of techniques based on the photoreceptor adaptation to transform pixel intensities into display pixel values have been introduced, as an example, Schlick's [10] quantization function F can be shown:

$$F(I) = \frac{pI}{pI - I + I_{max}} \tag{6}$$

where I is the pixel intensity, I_{max} is the maximum pixel value and p takes a value in the range $[1, \infty]$. Reinhard [7] proposes rewriting this function to:

$$F(I) = \frac{I}{I + \frac{I_{max} - I}{p}} \tag{7}$$

which can be then easily transformed into the aforementioned adapted photoreceptor equation by setting n to 1, and substituting $\frac{I_{max}-I}{p}$ for σ_b.

3 Prototype Camera Setup

This sub-clause provides an overview of the hard- and software, as well as the measurement environment used in this work. The prototype camera using state of the art HDR video technology, as well as the automotive software used to implement the IQ measurement criteria and framework (see 5) shall be described. The prototype vehicle setup enables the display and record of the aforementioned HDR video and will also be presented.

3.1 Image Sensor

Two details in particular set the applied HDR Complementary Metal Oxide Semiconductor (CMOS) image sensor hardware apart. Compared to the design used by other manufacturers, it does not require any additional image processing to create tone-mapped images. The image sensor itself is standalone and incorporates a full HDR tool chain including the capturing, construction and tone-reproduction of HDR images. Also, the image sensor is capable of simultaneously capturing two different exposures of a scene.

3.1.1 Image Sensor Architecture

Since the architecture of the image sensor is proprietary, some specifications and exact details are not freely available. However, a regular image sensor architecture can be complemented with the missing HDR components to highlight the process chain. The sensor contains three major blocks: the image sensor core, the image signal processor and the output interface. The image sensor core is responsible for collecting the electrical charge which is generated by the photo signal using the photodiodes. After that a readout of the accumulated charge is performed and the values are converted to a voltage signal for each pixel. Afterwards, the signal is amplified and converted to the digital domain using an analog-to-digital converter. The stream is than passed on to the image signal processor where the HDR processing is executed, which generates an output image for the target display. The output interface initiates a new readout and transports the already generated image via the LVDS interface to the next hardware component.

Pixel Design

The technique for an HDR image construction described in Sect. 2.1 requires capturing multiple exposures of the same scene successively and can therefore be referred to as temporal HDR. This approach would render the option of capturing a high frame rate video nearly impossible. The manufacturer has presented a unique pixel design in their image sensor, using two independent photodiodes in the active pixel array. Both of these photodiodes are able to capture the same image using different exposures simultaneously, henceforth, this approach can be referred to as spatial HDR. Motion artifacts are reduced when using this process. Since the capture of the exposures start at the same time, any movement in the scene is nearly not visible. This can be easily proven, given that t_s is the time required for a short exposure and t_l it the time required for a long exposure and those timings are equal for a spatial and temporal image acquisition, the following equation:

$$t_l - t_s < t_l + t_s \tag{8}$$

Shows that the time difference between a spatial and temporal HDR exposure process equals $t_s + (t_l - t_s)$.

It should be noted that the photodiodes are not only independent from each other in their exposure time, they also differ in size. The photodiode used for the short exposure is smaller than the photodiode used for the long exposure. Thus, the smaller photodiode induces the drawback that the sensitivity is reduced and therefore the noise level is increased. This aspect however, shall be discussed with timing and motion artifacts and sensitivity in mind. Reducing one photodiode's size frees up space for the other photodiode and makes it more sensitive to low light areas in the captured scene.

Additionally, the smaller photodiode is less sensitive to light and is therefore well suited to capture the bright areas of the scene, without overexposing. This results in two benefits: First dark scene regions can be captured with the highly sensitive large photodiode. The second benefit arises from the lower sensitivity of the smaller diode: to capture bright areas it is exposed with nearly the same exposure time as the large photodiode to capture the bright parts of the image, without overexposing. This results in similar motion artifacts in both exposures, which allows them to be merged in a single radiance map, without motion compensation. It is clear, that in the radiance map composition dark regions are covered by the large photo diodes, whereas bright regions are covered by the small photodiodes. Intermediate luminance values are combined by a ratio of both exposures. Some examples of scenes captured with the described HDR camera can be seen in Fig. 2. This architecture allows the image sensor to capture 30 frames-per-second HDR videos at a 1 megapixel resolution at two independent different exposures.

However, the exposure settings required for the weighting and composition of the radiance map are still to be tackled, as well as the subsequent tone-mapping procedure. This assignment is the main purpose of the next block of the image sensor, the signal processor.

Image Signal Processor

The image signal processor block consists of the components responsible for the creation of an image to be transported to the display using the output interface. The first step of the process as seen in Fig. 3 is the lens correction. The lens correction receives two separate data streams from the sensor core and handles both separately. The purpose of the lens correction routine is to minimize the vignetting effect created by the lens shading. This is achieved by applying an amplification on the corner pixels. The next step of the processing chain is the AWB component. The Color Temperature (CT) required for this procedure is estimated by the AWB routine itself. It should be noted, that up until this point the image signal processor has been working with two separate data streams handled individually.

The applied image sensor is overlaid with a color filter array, hence an interpolation is necessary to reconstruct the full color image from the camera. This process is generally known as demosaicing or debayering, named for the Bayer pattern used in the image sensors, where a two-by-two pixel area contains two green, one red and one blue pixels. Alternatively, the color interpolation of the image sensor can be switched off, and performed manually further in the toolchain. Using a state of the art linear demosaicing technique will also reduce the noise level [11]. However, the RAW image output necessary for this process cannot yet be

Fig. 2 From *left* to *right top* to *bottom* A rather well-lit twilight scene, note the captured illumination of the office buildings in the background. Next, a dusk scene with artificial lighting present, the lamps in the foreground are captured as well as the night sky in the background. Further, an artificially lit twilight scene, preserving dark details in the background. Lastly, a night scene with light spots produced by street lamps

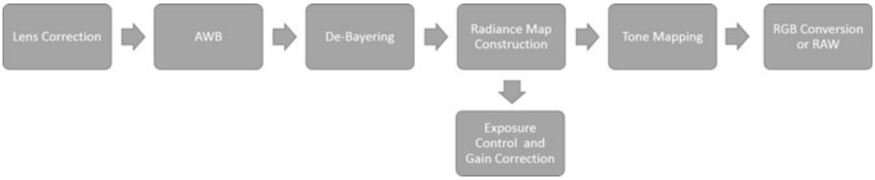


Fig. 3 The image signal processor pipeline with the complemented HDR components

provided by the image sensor at the same frame rate as an RGB output. In the next step, the HDR component performs the described de-bayering of the two 12bit streams. The radiance map is then composed, where the exposures are weighted by the response curve and exposure time. The result is one image stream containing a radiance map in an absolute photometric space covering up to 120 dB of absolute photometric values. With a known radiance map resolution of 20 bit, the mentioned dynamic range DR of 120 dB can be calculated as follows:

$$DR = 20\,log_{10}\,2^{20} = 120.41\,\text{dB} \tag{9}$$

This value coincides with the 120 dB dynamic range specified by the manufacturer. The HDR composition component also provides statistics and feedback to the exposure and gain control module, which in term calculates the exposure and gain settings for the next frame based on the current one. Tone-mapping is then applied on the radiance map to allow the image to be viewed on regular displays. The image can than either be transported as a raw stream or passed on to the color correction routine and then converted to the RGB color space.

3.2 Vehicle Prototype Setup

The prototype setup installed in the vehicle consists of three components. The already aforementioned HDR camera with the described image sensor which outputs a tone-mapped image to the Field Programmable Gate Array (FPGA) box used for de-serialization purposes, and the packaging of the stream into network packages and is connected via Gigabit Ethernet to the measurement system. The distinctive feature of this whole setup is the included backchannel functionality. The camera not only outputs an image and a current settings feed, but is also able to receive packages to reconfigure the image sensor in real-time without a usually necessary reboot of the system.

3.2.1 ADTF

The main purpose with which the ADTF was created was the rapidly growing amount of development work being sacrificed into solving portability and compatibility issues with different projects, software modules and hardware. It is not only the car of tomorrow, but also the car of today that is a very complex network of components (Controller Area Network bus, and later on the FlexRay just to name a few examples). Therefore, the need for a standardized software framework for vehicle based measurement systems to increase productivity during a development phase by enhancing the reusability and maintainability arose. ADTF is based on the module chain principle. An image processing module chain for example would initiate with the fetching of the image itself from a dedicated camera, if a live processing is assumed or any kind memory storage otherwise (physical storage, Random Access Memory (RAM) and so on). ADTF modules, or filters how they generally known, interact with each other through media samples of filters connected through pins which then transport the image frame to the next filter in the module chain. Depending on the task this process can of course be multiplied or even parallelized; the processed image would then either directly be displayed to the user, or be written to the aforementioned storage. The advantage which ADTF offers is the fact that the proposed image module chain can be changed at any time using provided or developed standardized ADTF components, a pipeline that simply flips or rotates an input image could then also sharpen the image, calculate a histogram and then output the image. While in the majority of cases this would require a complete redevelopment, in ADTF, the addition of two filters would suffice.

3.2.2 ADTF Toolchain

The standard tool chain for the utilized camera setup without the IQ modules shown in Fig. 4 consists of a configuration filter that allows live changes to the camera settings (e.g. exposure time, gain) which is connected to the grabber module responsible for the image retrieval from the camera. The grabber module has two output pins, the first one transports the packed image from the camera to the next block, and the second one outputs the camera's currently set configuration which is then being displayed on a generic ADTF table view component.

Due to bandwidth restrictions, the image from the camera is being transported in a packed format which uses reconfigured Ethernet packages and is able to arrange three pixels in two bytes of storage. This packed image format from the first output pin is then transported to the module which performs a de-packing and outputs the image as regular three channels RGB eight bit per channel stream. This stream can then be optionally cropped or rescaled or directly output to a generic ADTF display device. The already implemented standard components to read the Control Area Network bus can optionally be used to read the values of the car's twilight sensor or any other value present on the bus. The advantage of the framework is that the IQ

Fig. 4 The ADTF tool chain used to grab and display the HDR video stream from the camera and the according settings preferences module

assessment components can be seamlessly integrated into the given toolchain, without any modification whatsoever. The components are simply connected to the pin of the de-packaging module and can directly evaluate the output image in real-time.

4 Image Quality Issues

There are a number of IQ criteria mentioned before briefly: noise, color reproduction and blooming artifacts induced by point light. The exact causes for these issues shall be discussed in the following.

4.1 Noise

Noise occurrence itself can be divided into temporal and spatial noise. Temporal noises affect each pixel in the same way, while spatial noises result from differences between pixels. The further analysis shall however not divide noise appearance into temporal and spatial categories. Rather, a prospective from a CMOS image sensor pipeline point of view shall be taken. Gow et al. suggest a division into three blocks: charge collection and associated noise, pixel circuit noise and analog-to-digital conversion noise [12]. The charge collection is the first operation performed in the image sensor core block. The temporal noise which manifests itself during this process is known as photon shot noise. Dark current noise is introduced by image sensors reaction to temperature and manifests itself in both spatial and temporal noise, non uniformity and shot noise respectively. Pixel circuit noises consist of reset and thermal noises, both being temporal noises. Analog-to-digital conversion noise, which occurs when the signal values are being digitalized, is known as quantization noise.

4.1.1 Photon Shot Noise

As Hasinoff [13] points out, image sensors measure scene irradiance by counting the amount of light present on the sensor over a given time interval. In the physical domain, this light can be regarded as photons. The number of photons collected by a photodiode is uncertain due to their random individual arrival [13], or by extension the quantum nature of light itself [12]. This is called photon shot noise. In a given observation frame the photon shot noise is known to follow the Poisson distribution [12]. Photon shot noise can generally be reduced by taking multiple exposures of a scene and creating an average image.

4.1.2 Dark Current Noise

Even when light is not present on the image sensor, thermal energy can generate electrons which excite by the image sensor. This is known as dark current noise. Dark current noise is divided into non uniformity (spatial) and shot noise (temporal). Non uniformity dark current noise is a fixed-pattern exposure-dependent noise [12]. Dark current areas are distributed randomly and therefore cause different pixels to produce different dark current amounts. Dark current shot noise on the other hand is a temporal noise which occurs for the same thermal reasons, but affects each pixel the same way and is also, like the photon shot noise is known to follow a Poisson distribution [12]. Thus, Dark current noise is highly related to the image sensors temperature. To reduce this noise the power dissipation in the CMOS electronics of the sensor shall be kept to a minimum and the semiconductors shall be designed to minimize thermal noise. The applied image sensor described in 3 uses a process called black level calibration to deal with Dark current shot noise. The sensor has a certain amount of pixels which are not exposed to light, but are however processed the same way as the other pixels. The average values of the dark pixels is then subtracted from the pixels that are exposed to light.

4.1.3 Reset and Thermal Noise

The aim of the reset operation is the signal measurement of the electrons number generated from photons present on the photodiode. Reset noise originates from thermal noise causing voltage fluctuations in the standard reset level for a pixel [14]. As Calizo [14] points out, an already increased noise floor in low light conditions enhances the chances of a noisy reset operation, which in term decreases the pixel performance.

4.1.4 Quantization Noise

When the analog-to-digital converter creates a digital representation of the signal, rounding errors occur. This is called quantization noise. Increasing number of bits added to the image sensors analog-to-digital converter can reduce quantization noise.

4.2 Color Reproduction

Color genuineness and reproduction issues in the image can be caused by several reasons. Firstly, low-light conditions generally are not an ideal scenario, since contrast and color reproduction decrease with light deprivation. Secondly, the applied image sensor tends to turn down the saturation when sensing low-light conditions. This is done to decrease color noise, which is much more distracting than black and white noise, however, this process introduces a noticeable desaturation effect. Thirdly, a new kind of image artifact is produced by HDR imaging. So called grey artifacts are present in some images, when the middle range of the color spectrum is not being covered well enough by the exposures taken to calculate the radiance map.

4.3 The Point Light Sources Issue

Blooming and lens flare artifacts from headlights in low-light and dark scenes, although being a major IQ issue, can be explained quite easily. When capturing images in low-light conditions, the image sensors exposure is generally set to the highest possible value, in the case of video capturing, the highest possible value that does not affect the given frame rate. When however a bright light source appears in the scene, the set exposure is generally far too high to capture the light source accordingly, causing the amount of light which hits the photodiode to "overflow" to neighboring photodiodes. Blooming effects are more prominent in charge-coupled device (CCD) sensors, while CMOS image sensors suffer more from lens flare artifacts, which are introduced by reflection off lens irregularities. Figure 5 demonstrates an example of point light sources blooming and lens flare artifacts. In dark scenes, the long exposure is favored by the image sensor, since it contains less noise and is assumed to have more details of the scene. Point light sources can be dealt with by changing the image sensors automatic exposure control and first creating the short exposure which captures the bright light spots of the scene.

Fig. 5 From *left* to *right*, the first image shows distinguishable headlights at a distance of 100 m, while the second image shows the headlights at a distance of 300 m, which are not separated anymore due to high blooming and lens flare

Second, when composing the radiance map, the weighting must be corrected to prefer the short exposure for the point lights, since those are correctly exposed, compared to the long exposure which contains the artifacts.

5 Image Quality Criteria

Automotive driver assistance systems nowadays are becoming more and more complex, ranging from highly intelligent braking systems to camera systems being capable of adjusting the cars lane, or recognizing traffic signs [15]. The rapidly growing range of camera based assistance systems utilizing CMOS sensors, has proven itself very challenging in terms of measuring and later on, guaranteeing its image quality [16]. As Hertel [16] points out, typical camera specifications such as resolution and frame rate might be helpful when assessing the devices in a laboratory conditions. However, automotive scenes include a wide range of different conditions such as very sparse light, low and high beams in the dark and sunshine on the other hand. Wang et al. [17] remark that it is essential to develop objective image quality assessment measures which values align with perceived image quality.

Objective quality metrics can be divided into three categories according to the availability of a distortion free reference image, also known as a ground truth [17]. A full-reference assessment assumes a complete original reference image. A reduced-reference quality measurement is performed with a partially available ground-truth image, while a no-reference, also known as “blind” is performed with no distortion free image at all [17]. This work is based entirely on full-reference image quality assessment and introduces a ground-truth tool, which is capable of calculating a nearly perfect distortion free image for any given scene in the matter of several hundred frames. Hertel [16] proposes measuring the cameras sharpness, image noise level and tone reproduction. Since in this particular case the image is being directly presented to the driver, Xie et al.’s [18] toolset of image quality measurement techniques based on the HVS are also introduced, which consist of the average luminance of an image, average entropy information of an image, and average contrast of an image. The test-methods of the newly drafted ISO 16505 [19] standard are also used in the case of point light sources measurement and optimization. A tool to evaluate the cameras performance with low beams in the scene is introduced.

5.1 Ground Truth of an Image

The most efficient way to calculate a noise free reference image of a scene can be achieved by calculating the sum of several frames and dividing each pixel value by the number of frames used to calculate the sum. This must be done under the assumption that all of the frames used are equal in respect to all camera settings. The sum for a given pixel (i,j) is then calculated as follows:

$$S_{ij} = \frac{\sum_1^f I_{R,G,B}}{f} \quad (10)$$

where $I_{R,G,B}$ stands for the red, green and blue values of the pixel at location (i,j) respectively and f is the number of frames used to calculate the sum.

5.2 Average Entropy

Shannon [20] introduced the concept of entropy in 1948. First, for simplicity, let us consider an 8 bit gray image with 256 possible intensity values. The Shannon entropy is then defined as follows:

$$E_S = -\sum_1^{256} p_i \, log_{10} \, p_i \quad (11)$$

where p_i is the normalized frequency for the possible intensity levels. Basically, as Hou et al. point out, the function measures the uncertainty of a pixel to take an intensity value from the possible number of gray levels. A uniform image histogram would produce a high entropy value, whereas a histogram of an over- or under-exposed image would decrease the information amount, and by that, the entropy. Empirically measured, the average entropy values in this work ranged from approximately 3.0 for a dark scene to 5.0 for a well-lit twilight scene. Since the average entropy calculation prefers uncertainties in the scene, which image noise is, it is calculated over the ground truth image. A calculation of the entropy using the noisy image would mistake the noise for actual scene information and produce an incorrect value. For a better understanding of the information content in an image, a 10 × 10 region entropy calculation, which visualizes the values on a color scale was introduced. A direct visualization of the information in a given scene makes the calculated entropy value far more comprehensible. An example can be seen in Fig. 9. This image also shows why the ground truth calculation is absolutely necessary. The homogeneous concrete in front of the car, which contains a lot of noise, is rewarded by the entropy calculation and is said to contain the same information as the car's body, which it does not. This would lead to incorrect assessment of the visualization of the scene, but also increase the average entropy calculations value. Using the ground truth image, only the actual "real" information in the scene can be measured.

5.3 Peak Signal-to-Noise Ratio

While there are multiple components and causes for noise occurrence in image, the calculation of the noise level in the resulting image is the same for all types of noise. The idea behind the Peak Signal-to-Noise Ratio (PSNR) calculation is to subtract the noisy image from the distortion free reference image. The higher the PSNR value means, the less noise is in a scene, with values for 8bit per pixel images ranging from 25 to 40 dB [21]. However, the calculation of the PSNR by its own cannot be regarded as a criterion. A completely black image without any noise, would produce a high PSNR value, but would not contain any valuable information. The average entropy value on the other hand, would be very low, since uncertainty level of another value rather than black would be low. Hence, the PSNR value and the average entropy values must be regarded together for a correct assessment of a scene. A PSNR calculation is performed as follows, first, the Mean Squared Error (MSE) must be obtained:

$$MSE = \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}(S(i,j) - N(i,j))^2 \tag{12}$$

where m and n are the width and height respectively, S is the distortion free summed pixel and N is the regular, noisy pixel of a single frame.

The PSNR is then defined as follows:

$$PSNR = 20\,log_{10}\frac{I_{max}}{\sqrt{MSE}}\,dB \tag{13}$$

where I_{max} is the maximum possible intensity of the signal, which in case of an 8-bit image is 255.

5.4 Average Luminance

Xie and Wang [18] propose calculating the average luminance AL using the following equation:

$$AL = \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}L(i,j) \tag{14}$$

where m and n are again the width and height of the image, and $L(i,j)$ represents the intensity of a pixel at location (i,j). When calculated in the HDR domain, this value can be regarded as the representation of the amount of light in the scene. Since however, the radiance map is tone-mapped to the LDR domain, the average luminance calculated for a tone-mapped image cannot be viewed as an absolutely trustworthy depiction of the scene luminance. Therefore the measurement of this particular criterion must be complemented by a physical value not dependent on any camera specific settings. In this case, a photometer positioned right above the camera was used to capture the amount of light in lux falling on the camera lens, or by extension, the image sensor itself.

5.5 Average Contrast

Contrast can be defined as a basic perceptual attribute of an image, and while many studies in the field of visual perception have gone into contrast sensitivity and contrast enhancement, there are still no clear definition on the measurement of contrast itself. The Michelson [22] contrast $C_{\mathrm{Michelson}}$ is defined as follows:

$$C_{Michelson} = -\frac{I_{max} - I_{min}}{I_{max} + I_{min}} \tag{15}$$

where I_{max} and I_{min} are the maximum and minimum intensity values, respectively.
The Weber [22] contrast C is expressed as:

$$C_{Weber} = \frac{\Delta I}{I} \tag{16}$$

where ΔI is the increment or decrement in the target intensity from the uniform background intensity I. While the calculation of the contrast using both of the mentioned approaches does not share any similarity in the value range [22], they do however both represent contrast as a ratio of luminance change to mean background luminance. Also, those calculations assume a single target and a uniform background, which is not the case in this work. The calculation using the root-mean-square contrast is far more applicable, since it is used for image comparison [22] and does not require a uniform background or a single target.

$$C_{RMS} = \sqrt{\frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}(I(i,j) - I_A)^2} \tag{17}$$

where m and n are the width and height of the image, $I(i,j)$ represents the intensity of a pixel at location (i,j) and I_A is the overall average intensity. The intensities of the images are assumed in a normalized range $(0, 1)$. Basically, a calculation of the separation between the darkest and brightest areas in the image is performed, with a higher value representing a better contrast.

5.6 Color Rendering

The color rendering measurement in this work was performed using the original x-rite colorchecker colors of the Digital SG colorchecker. The benchmark values provided by x-rite were used to calculate the color deviation in any given image.

Due to the already mentioned desaturation and grey artifacts introduced by the image sensor, the color values in the tone-mapped image will not perfectly align with the benchmark values of the color checker, therefore, the error rate must be measured. The colors red, green, yellow and white in this case are of higher significance, since those correspond to traffic lights, head and brake lights. Since, unlike the other calculations, those values could not be provided in real-time. The Average Absolute Color Deviation from Greyscale (AACDG) was introduced to provide real-time feedback towards the image's overall "colorfulness". Similar to the contrast calculation, the AACDG measures the spread of the color values in the image compared to their greyscale representation, with a higher value meaning more colors in the image. Hence, an image with grey artifacts would score lower on this scale. The AACDG is calculated as follows:

$$AACDG = \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}\left|(AI_{R,G,B}(i,j) - I_G(i,j))\right| \tag{18}$$

where m and n are the width and height of the image, $AI_{R,G,B}(i,j)$ represents the average color intensity of a pixel at location (i,j) and $I_G(i,j)$ the grayscale intensity of the same pixel.

5.7 Point Light Sources Measurement

To make the measurement of the blooming artifacts caused by point light sources in the scene possible, a tool capable of plotting a single image line in real-time was introduced. The plot would include the pixel count on the x-axis, and the pixel intensity on the y-axis. This values would directly show the distinguishability between two point light sources and the pixel values intensity of dark and light areas ratio in every possible scenario. Since such measurements must be performed in complete darkness, with a distance between the light sources of approximately 1 m (average distance between automobile head lights) and a distance range to the camera of 100–300 m, a Light-Emitting Diode (LED) panel with scaled measurements seen in Fig. 6 was assembled.

Distinguishability of the scene by the driver is of high importance, hence so is the separation of the point light sources. When looking at Fig. 5, one cannot assess neither the distance to the car, nor even if it is in fact a car. The single light source might be mistaken for a motorcycle or even a bicycle.

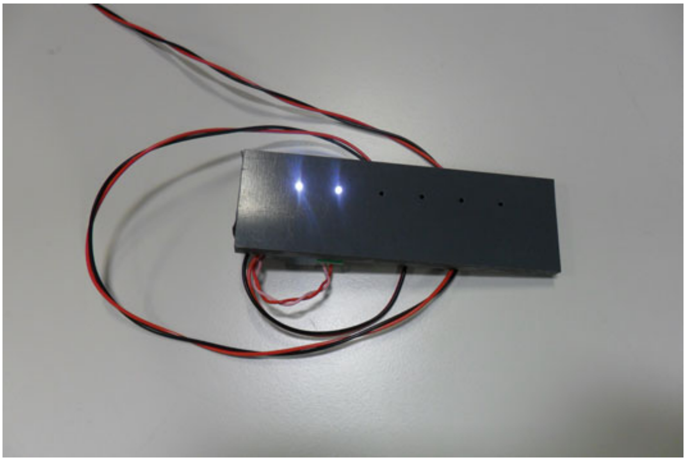


Fig. 6 The current positioning simulates a head light distance of 0.7 at 300 m, further adjustment of one LED to the right would allow distances of 0.9/1.1/1.3/1.5 and 1.7 m respectively

5.8 Sharpness

Since the camera system in this automotive application is mounted with constant focus lens, the sharpness itself does not change over time, and must therefore, be measured only once. The Modulation Transfer Function (MTF) is the comparison of the edge contrast of an object to its image representation. An MTF measurement determines the point at which the camera's sharpness reaches a certain drop. An MTF90 calculation would determine the point, at which the sharpness of an image decreases by 10 % from its highest point is measured. To compensate for image sharpening algorithms used by the image sensor, an MTF90P measurement is performed. In this case, the values are firstly decreased by 10 % to account for the intensity peaks introduced by the image sensor. Such peaks can clearly be seen in Fig. 7.

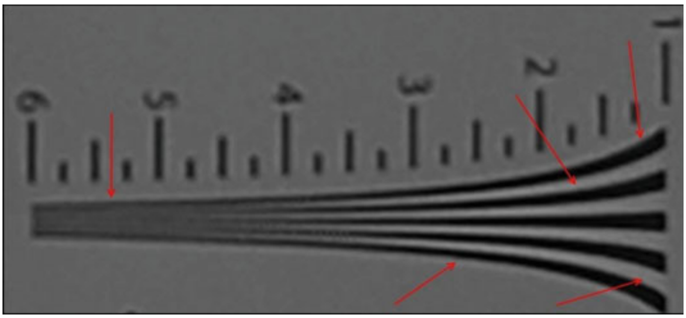


Fig. 7 Intensity peaks around the dark edges (pointed out by the *red arrows*) introduced by the sharpening algorithms of the image sensor

An MTF90P measurement using the ISO12233 test chart, which is used in the ISO16505 [19], is performed as follows:

1. An image of the test chart with the maximum size possible is captured using the camera at a defined distance. Since the utilized camera is focused to infinity and performs equally after a certain distance, in this case 1 m, this distance was chosen.
2. Single lines of the image with the so called wedges are evaluated according to their greyscale values.
3. Once the gray level values of the lines reaches the previously mentioned 10 % decrease, the corresponding number which stands for the lines ×100 per picture height represents the overall sharpness of the camera, where higher values represent a sharper image.

5.8.1 Exemplary Results

Since the MTF measurement was only performed once, due to the fixed aperture of the applied camera, the image sharpness did not change over the course of this work and was not optimized. The measurement however was performed by positioning the wedges in the middle and corners of the image. This was done to determine the sharpness drop introduced by the form of the camera lens. The results of the measurement can be seen in Table 1, which as predicted, shows a slight sharpness drop in the corners of the image.

5.9 Examples

Figure 8 shows the results of a reference image calculation, and how the use of the wrong image as a reference can lead to corrupt results.

The region entropy calculation depicts how the average entropy calculation would reward noisy pixels and is therefore by itself is not a good image quality criterion.

Table 2 shows an example calculation of all the aforementioned criteria for the scene in Fig. 8. The resulted values are comprehensible: the very low amount of light, evident by the luminance value and the value measured by the photometer, leaves some parts of the image completely dark. The rather high PSNR value is also allegeable, although some noise can clearly be seen (e.g. on the white plan surface

Table 1 Results of the sharpness measurement

Location	According value
Center	0.516 LW/PH
Corners	0.45 LW/PH

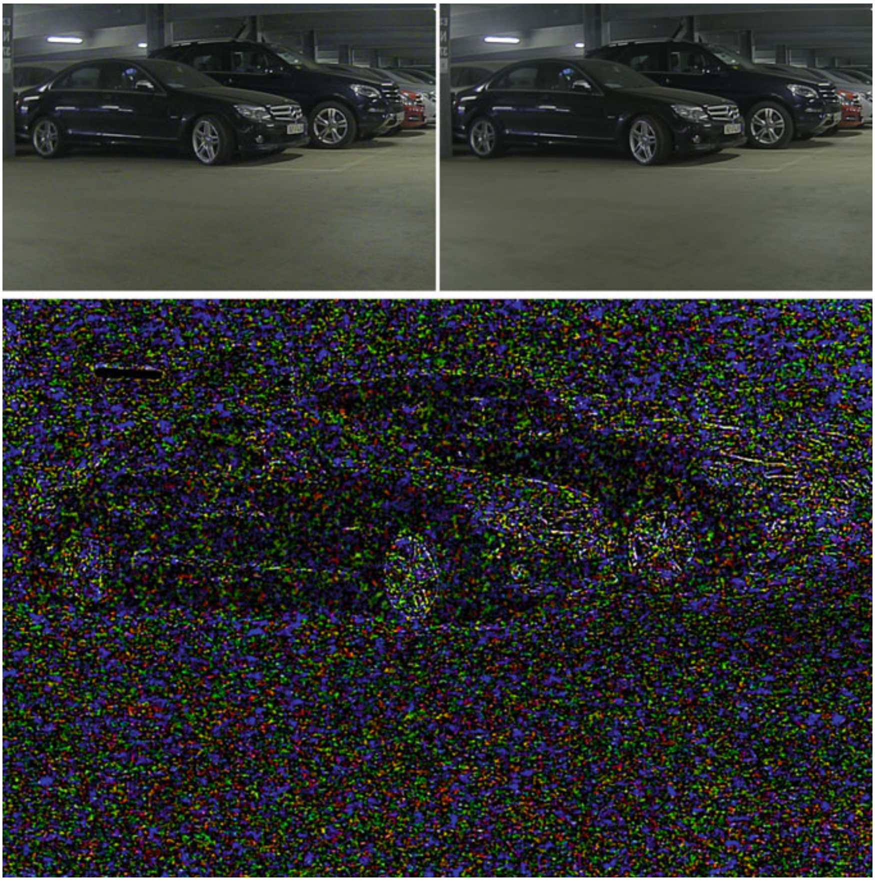


Fig. 8 From *left* to *right top* to *bottom*, the noisy original image of the scene, followed by the calculated reference ground truth image, and finally a subtraction of the images from one another to show the high noise level present in the scene

in the background) the aforementioned low light conditions leave big homogeneous dark spots without any noise at all.

The low contrast and AACDG values are to be expected, the former due to the general homogeneity of dark areas in the scene, and their resemblance to the background, the latter due to the small amount of colors, both those reasons also explain the low average entropy value in the scene. Once again the comparison of the PSNR value against the average entropy must be mentioned. The image provides a small amount of information at a low noise level, however, when presented with the decision, one should always opt for the higher entropy and by that information level rather than a lower noise level (Fig. 9).

Table 2 Results and camera settings for Fig. 8

Criterion	According value
PSNR	29.66 dB
Entropy	3.604
Luminance	17.247
Contrast	0.107
AACDG	1.88
Photometer at camera	0.13 lux
Camera setting	Value
Long exposure	33 ms
Short exposure	33 ms
Long gain	250
Short gain	50

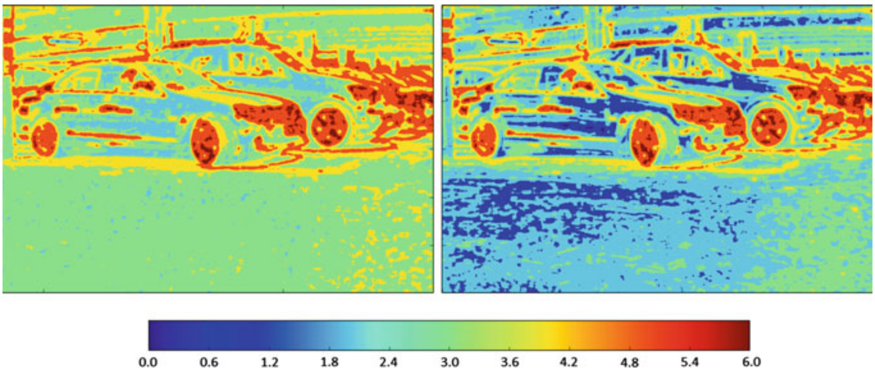


Fig. 9 A 10 × 10 region entropy calculation of a ground truth image (*right*), and a noisy image (*left*), where a higher value is interpreted as more information in the image. The 10 × 10 entropy region calculation clearly shows how a distorted image (*left*) would mimic a false high information level

6 Evaluation and Results

6.1 Optimization Approaches

One of the main disadvantages of the image sensor is that it is very easily forced into extreme low-light settings, in which, all of the exposure and gain settings are operating at maximum level. While this is useful in dark conditions and is beneficial for the color reproduction, twilight and point light sources scene suffer a quality degradation under these settings. This offers room for improvement. By manually adjusting the image sensor settings, an analysis shall show which automatisms can be further tuned to increase the low-light, twilight and point light sources performance. Using the developed real-time ADTF modules of the defined IQ assessment set, various optimization can directly be performed and evaluated.

6.1.1 Noise and Color Rendering

Image noise level reduction generally tends to only be motivated by the PSNR value itself, and does not take the image information, color or luminance into account [21, 23–25]. This approach does not consider the trade-off which then occurs when a PSNR increase is achieved simply by performing gain adjustment procedures for example. While the PSNR value will of course increase, a gain decrease will lead to an information and luminance loss at the same time. Especially in automotive scenes, such an approach shall not be considered, since the image information is far more significant. While a lower noise level is undoubtedly an advantage, a pedestrian (see Fig. 10) in the scene lost as the consequence of a gain reduction is not. Therefore the optimization of the noise level shall proceed under a constant monitoring of the overall information content depiction of the scene in the image. Possible image sensor components to be evaluated in this case are the AWB gain level, lens correction gain level and image sensor de-noise level. In parallel to that, it shall be evaluated which gamma and saturation levels can be set to increase the color rendering in the image. Also, the exposure settings ratio of the image sensor must be optimized to avoid grey artifacts in the scene.

Fig. 10 Two images at exactly the same conditions but with different long gain settings, while the pedestrian is barely visible in the *left* image, he is still distinguishable in the *right* image

6.1.2 Point Light Sources Optimization Techniques

The problem of the point light sources in the image is created when light spots are introduced into a low-light scene. The exposure times of the image sensor in this case are not suitable to capture the light sources accordingly. Yet, an analysis of the RAW images from the sensor showed, that when manually decreasing the short exposure, the light sources are clearly distinguishable. However, further processing in the pipeline heavily favors the long exposure and the blooming is introduced. Thus, dedicated image sensor settings different from the regular low-light settings are necessary to correctly capture point light sources in a scene.

Manual sensor settings adjustment have shown, that in several cases, especially when dealing with twilight and point light sources scene, and improvement of the IQ can be achieved. The behavior of the manual settings and the resulting automatic settings as well as the current image sensor settings must now be shown and evaluated.

6.2 Manual Camera Settings Results and Optimization

6.2.1 Linear Light Increase

In laboratory conditions, the light amount was gradually increased from 0.03 lux to 1.05 at the camera in 20 steps and the scene recorded for evaluation. The camera settings which were

constant the entire time can be seen in Table 3, the results can be seen in Table 4. Several conclusions can be drawn from the measured values and already made assumptions also proven. It can clearly be seen, that PSNR values worsen by 19 % over time during a light increase in a scene where a light shift from very dark to twilight conditions is performed. The information in the image shown by the entropy values increases by 57 % over time, as well as contrast and color rendering shown by the AACDG and the percentage of achieved colors in comparison to the original colorchecker colors increases. Also, the almost perfectly linear light response function of the HDR camera can be concluded. It can be drawn that all criteria except for the PSNR values increase when more light is introduced into a scene.

Table 3 Camera settings for the linear light increase measurement

Camera setting	Value
Long exposure	33 ms
Short exposure	33 ms
Long gain	255
Short gain	64

Table 4 Linear light increase from 0.03 lux to 1.0 at camera

Lux	PSNR (dB)	Luminance	Entropy	Contrast	AACDG	Color (%)
0.03	32.07	9.02	2.88	0.065	1.067	23
0.1	28.74	17.09	3.6	0.11	1.63	30
0.15	28.2	19.55	3.72	0.12	1.91	52
0.2	28	20.86	3.82	0.132	2.37	58
0.25	27.9	22.65	3.9	0.136	2.62	60
0.3	27.7	22.88	3.949	0.139	2.78	65
0.35	27.6	24.5	4.02	0.141	2.83	66
0.4	27.7	26.57	4.09	0.145	2.91	68
0.45	27.57	28.29	4.15	0.148	3.06	69
0.5	27.4	29.77	4.2	0.151	3.274	73
0.55	27.2	31.03	4.25	0.154	3.655	74
0.6	27.1	32.67	4.31	0.157	3.87	77
0.65	27	34.25	4.34	0.159	4.01	79
0.7	26.8	35.51	4.377	0.161	4.25	80
0.75	26.7	36.96	4.4	0.163	4.56	82
0.8	26.6	37.9	4.43	0.165	4.7	85
0.85	26.3	39	4.45	0.168	4.89	86
0.9	26.5	40.45	4.48	0.17	5.16	87
0.95	26.2	43.194	4.53	0.18	5.59	89
1	26	44.05	4.55	0.182	5.77	91

The drop in the PSNR values can be explained by the fact that while the light amount is greater than before, the shift is not performed towards very bright conditions which can be compared to a sunny day, but rather twilight conditions. Hence, while the light falls on and by that adds more information and color to the scene, homogeneous dark regions that were not seen before are now visible and subject to noise occurrence.

6.2.2 Linear Gain Increase

In laboratory conditions, the image sensor configuration remained constant for all but one setting: the gain value of the long channel, which was increased from minimum to maximum setting in 9 steps. When sensing dark conditions, the long exposure is weighted higher by the image sensor during the radiance map composition, therefore the long channel gain increase can be directly retraced in the output image. The image sensors negligence of the short exposure is explained by the fact that smaller photodiodes produce a higher noise level [1], as well as the far smaller maximum gain increase possible (250 for the long exposure versus 60 for the short exposure). The camera settings for this sequence can be found in Table 5 and the results in Table 6. This results show the already stated assumption that the

Table 5 Camera settings for a linear gain increase from 50 to 250

Camera setting	Value
Long exposure	33 ms
Short exposure	33 ms
Long gain	50–250
Short gain	50
Photometer at camera	0.13 lux

Table 6 Linear gain increase from 50 to 250

Gain	PSNR (dB)	Luminance	Entropy	Contrast	AACDG
50	32.01	6.16	2.57	0.04	0.94
75	31.26	6.94	2.62	0.05	1.14
100	30.05	8.31	2.8	0.06	1.4
125	29.56	10.22	3.05	0.07	1.64
150	29.3	10.83	2.92	0.08	1.7
175	29.2	13.52	3.35	0.09	1.95
200	29.2	14.75	3.44	0.1	1.92
225	29.1	15.62	3.49	0.11	1.9
250	29	17.24	3.6	0.12	1.9

information content in a scene increases with higher gain settings. While the PSNR values again have clearly dropped by 10 %, a 40 % information increase over time can be seen.

6.3 *Automatic Camera Settings Results and Optimization*

6.3.1 Noise and Color Rendering

In laboratory conditions, a scene was first captured with the automatic image sensor settings. The PSNR level of the scene was used as a baseline, while the sensor automatisms were changed in an effort to increase information content, color reproduction, contrast and the AACDG. The effort in this process was to produce an increase of the PSNR level, which in terms can be used to boost the gain settings. The goal was not to achieve a lower noise level, but rather an information increase in the scene at constant PSNR baseline. Several settings were used for this approach. First, the digital image sensor was decreased and the lens correction gain was turned off, this offered for two analog gain increases. Also, the de-noise function was altered to concentrate more on the short exposure, since the noise level there is higher due to the smaller pixel size. This once again allowed for another gain increase.

6.3.2 Point Light Sources Optimization

As already mentioned, the short exposure in the capturing process does contain the accurately captured point light sources in a scene. A new configuration was introduced for the tone-mapped image to display the point light sources correctly to the driver. First, a new exposure ratio was set, forcing the image sensor not to increase the short exposure in dark conditions. Second, the weighting procedure was altered in this case, changing the radiance map composition. The short

Table 7 Camera settings that made the correct capture of the point light sources possible

Camera setting	Value
Automatic settings	
Long exposure	33 ms
Short exposure	33 ms
Automatic exposure control	Standard
Exposure weighting	Standard
Changed settings set	
Long exposure	33 ms
Short exposure	4.1 ms
Automatic exposure control	1/8 ratio short/long
Exposure weighting	80/20 % ration of short/long

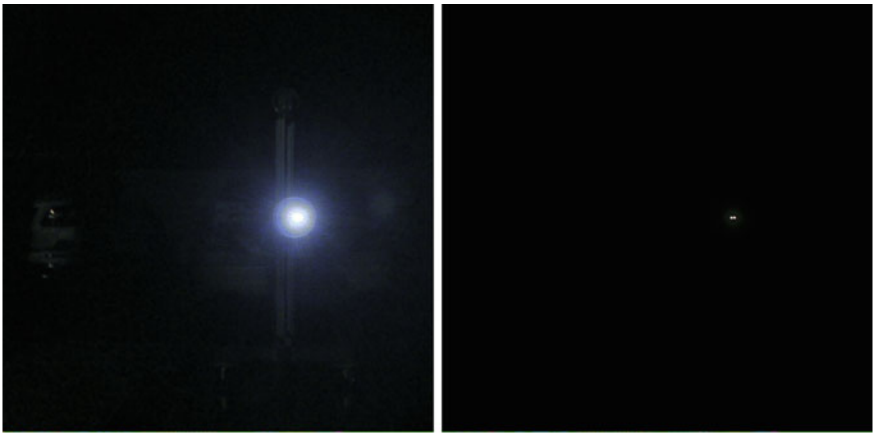


Fig. 11 The first image (*left*) shows the standard image sensor settings when capturing point light sources, while the second image (*right*) shows the optimized set where the point light sources are clearly distinguishable. It should be noted however, that the image quality itself may decrease, and field test study must determine which is of the settings is better for the visual perception

exposure would now be preferred over the long exposure, leading to a correct output of the point light sources. The results can be seen in Fig. 11, and the camera settings are listed in Table 7.

7 Conclusion

This work first discussed dynamic range in the imaging context. The advantages of HDR imaging compared to regular LDR imaging were presented. The processes of capturing, calculating and then displaying HDR radiance maps were then shown. This however, opened room for the next issue: How, in terms of hardware, can the problem of HDR imaging, namely, the capture and display be solved fast enough to be displayed as a tone-mapped high frame rate video.

Further, the complete hard- and software toolchain required for such a task was presented. The prototype setup, as well as the automotive framework used for the development of the real-time modules of the IQ criteria was described.

Subsequently, the issues of the applied prototype HDR camera were elaborated. While some, such as the grey artifacts, are introduced by HDR imaging, noise and color reproduction are known problems that already exist for LDR cameras. All of these, however, must be clearly measured for any kind of optimization for providing a better result.

A set of criteria to objectively measure the IQ of a given scene was defined. The full- reduced and blind-reference image assessment techniques were described. The proposed ground truth approach allowing a full-reference scene evaluation using the described IQ criteria representing information content, color, noise level and luminance was discussed.

It was shown that twilight and low-light scenes must be handled differently. While a light increase at daytime will decrease the noise level, in low-light conditions, this is evidently not the case. Hence, a gain value which leads to the highest possible information content in very dark scenes and an acceptable noise level in twilight scene was defined. These examples also prove the established approach for the noise level reduction.

The applied image sensor, which uses two simultaneous exposures to capture HDR radiance map and output a high framerate tone-mapped video while already showing high low-light sensitivity, still struggles with dark scenes. Current state of the art image sensors offer up to three exposures, but captured as temporal HDR. An evaluation and comparison of such an image sensor to the one applied in this work would show which approach (temporal or spatial HDR) is more suitable for automotive low-light application. An evaluation of other IQ criteria could also provide even better understanding of a scene. Hou et al. [26] introduced Visible Entropy, which in experimental results, has been shown to correctly evaluate and compare HDR images without the need for a full-reference image, providing the precise visible information. A real-time implementation could prove itself to be very useful since it would be able to instantly feed statistics to a feedback loop and optimize the camera without any delay caused by the ground truth image calculation.

References

1. Reinhard, E., Heidrich, W., Debevec, P., et al.: High Dynamic Range Imaging: Acquisition, Display, and Image-Based Lighting. Amsterdam, Morgan Kaufmann (2010)
2. DiCarlo, J., Wandell, B.: Rendering high dynamic range images. Electron. Imaging **3965**, 392–401 (2000)
3. Schulz, S., Grimm, M., Grigat, R.: Optimum Auto Exposure Based on High-Dynamic-Range Histogram, pp. 269–274 (2007)
4. Chaurasiya, R., Ramakrishnan, K.: High Dynamic Range Imaging, pp. 83–89 (2013)
5. Birchfield, S.: An Introduction to Projective Geometry (for Computer Vision). Unpublished note, Stanford university (1998)
6. Kang, S., Uyttendaele, M., Winder, S., Szeliski, R.: High dynamic range video. ACM Trans. Graph. (TOG) **22**, 319–325 (2003)
7. Debevec, P., Malik, J.: Recovering high dynamic range radiance maps from photographs. In: ACM SIGGRAPH 2008 Classes 31 (2008)
8. Mitsunaga, T., Nayar, S.: Radiometric self calibration (1999)
9. Mann, S., Picard, R.: On being 'undigital' with digital cameras: extending dynamic range by combining differently exposed pictures, 7 pp. (2015)
10. Schlick, C.: Quantization techniques for visualization of high dynamic range pictures. In: Photorealistic Rendering Techniques, pp. 7–20 (1995)
11. Malvar, H., He, L., Cutler, R.: High-Quality Linear Interpolation for Demosaicing of Bayer-Patterned Color Images, 485 pp. (2004)
12. Gow, R., Renshaw, D., Findlater, K., et al.: A comprehensive tool for modeling CMOS image-sensor-noise performance. IEEE Trans. Electron Devices **54**, 1321–1329 (2007)
13. Hasinoff, S.: Photon, poisson noise. Comput. Vis. **35,** 608–610 (2014)
14. Calizo, I.: Reset noise in CMOS image sensors (2005)
15. Fleming, W.: New automotive sensors—a review. Sens. J. IEEE **8**, 1900–1921 (2008)
16. Hertel, D., Chang, E.: Image Quality Standards in Automotive Vision Applications, pp. 404–409 (2007)
17. Wang, Z., Bovik, A., Sheikh, H., Simoncelli, E.: Image quality assessment: from error visibility to structural similarity. IEEE Trans. Image Process. **13**, 600–612 (2004)
18. Xie, Z., Wang, Z.: Color Image Quality Assessment Based on Image Quality Parameters Perceived by Human Vision System, pp. 1–4 (2010)
19. ISO/FDIS 16505:2014(E): Road vehicles—ergonomic and performance aspects of camera-monitor systems—requirements and test procedures, International Organization for Standardization (2014)
20. Shannon, C.: A mathematical theory of communication. ACM SIGMOBILE Mob. Comput. Commun. Rev. **5**, 3–55 (2001)
21. Huynh-Thu, Q., Ghanbari, M.: Scope of validity of PSNR in image/video quality assessment. Electron. Lett. **44**, 800–801 (2008)
22. Peli, E.: Contrast in complex images. JOSA A **7**, 2032–2040 (1990)
23. Yao, W., Li, Z., Rahardja, S.: Noise Reduction for Differently Exposed Images, pp. 917–920 (2012)
24. Lee, S., Ortega, A.: A Novel Approach of Image Compression in Digital Cameras with a Bayer Color Filter Array, pp. 482–485 (2001)
25. Ponomarenko, N., Battisti, F., Egiazarian, K., et al.: Metrics performance comparison for color image database (2009)
26. Hou, Z., Yau, W.: Visible Entropy: A Measure for Image Visibility, pp. 4448–4451 (2010)

Intuitive Motion and Depth Visualization for Rear-View Camera Applications

Christoph Rößing

Abstract Human perception of depth and motion is strongly influenced by pictorial cues. These cues affect very basic neuronal processing structures, and lead to a modulated perception of distance and velocity on a higher perceptual level. Thus, it is proposed to convey a previously acquired distance and velocity information by manipulating or introducing motion and monocular depth cues in images and video streams to support human visual scene assessment. The addressed techniques utilize artificial depth of field, exaggerated motion blur, and color coded risk potential renderings. Except of the latter, these renderings are designed to maintain the natural look of the input material as much as possible. As a result, the supplemental distance and velocity information is conveyed to the viewer in a natural but distinct way. Previous to the rendering of depth and velocity cues in images and video material, this information has to be acquired first. Therefore, techniques for the scene reconstruction are handled in the first section, followed by the description of the newly presented rendering methods.

Keyword Perceptual rendering · Monocular depth cue rendering · Artificial motion blur · Depth-of-field rendering

Disclaimer The following chapter is part of *Video and Image Manipulation for Enhanced Perception* from C. Rößing [1].

C. Rößing (✉)
Department of Computer Graphics, Eberhard Karls University, 72076 Tübingen, Germany
e-mail: christoph.roessing@gmx.de

A. Terzis (ed.), *Handbook of Camera Monitor Systems*,
Augmented Vision and Reality 5, DOI 10.1007/978-3-319-29611-1_15

1 Monocular Spatial and Temporal Scene Reconstruction on Highways

To convey distance information via pictorial depth cues, depth maps have to be generated first. Therefore stereo camera based algorithms which implement dense depth map reconstruction methods (e.g. Semi Global Matching [2]) may be utilized.

To convey object trajectories and velocities in still images and video streams, a dense map of the object trajectories between consecutive video frames is required. In computer vision, these dense trajectories fields are referred to as dense optical flow [3].

Accordingly dense stereo and optical flow methods are powerful computer vision algorithms for the general purpose of spatial and temporal scene reconstruction. However, they do not provide a classification of the visible objects in the scene. For applications in which it is crucial to depict or emphasize, the motion or distance of a specific object a different approach has to be taken.

Thus, the generic computer vision tools of dense depth and optical flow reconstruction are replaced by an optimized method for highway traffic scene reconstruction: A monocular rear-view vehicle mounted camera setup is utilized to estimate the road course, the vehicle distances, and their trajectories. This output data is post-processed in a perceptual motivated rendering stage to enrich the rear-view camera video stream with motion and distance cues.

As a result, a specialized solution for supporting drivers during their assessment of the rearward traffic on highways is proposed: The camera feed that is used for scene analysis is manipulated in the rendering stage to depict motion, distance, and risk potential of trailing vehicles (see Sect. 4) and displayed afterwards to the driver.

In this section, a monocular camera sensor based approach for motion and distance estimates of trailing vehicles as well as road course reconstruction is proposed. The vehicle mounted sensor setup is shown in Fig. 1. A forward-looking camera is utilized to reconstruct the road course ahead whereas the odometry sensors are used to propagate this previously captured lane information behind the vehicle.

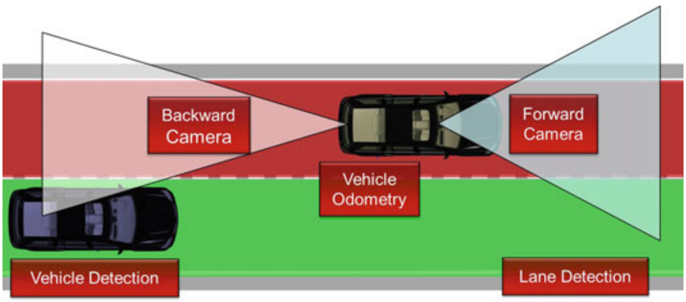


Fig. 1 Vehicle Sensor setup: A forward-looking camera is detecting the road course which is translated in accordance to the estimated ego motion and projected back into the backward camera image. Subsequent, the trailing vehicles are detected and merged into the combined environment model *M*

The backward-looking camera image is initially fed into a vehicle detector to segment the trailing vehicle contours and positions inside the image. To build up a combined environment model $M = [s_l, s_d, \ldots]$ which unites the road course s_l and the position sensor data of the vehicles s_d, the output of both detection algorithms is translated and rotated according to the ego-motion estimation. This model is updated in each time-step t to keep it in sync with the current vehicle surrounding.

It shall be noted that the utilized approach does imply a flat world assumption which clearly does not hold for all real-world scenes. However, for highway reconstruction scenarios, this is a commonly made assertion. If an extension to 3D models is required, the proposed approach can be extended with the methods introduced in [4].

The environmental model provides all information that is required for a subsequent rendering to enrich the image with perceptual motivated depictions of motion and distance. These renderings aim to subconsciously influence the driver's perception to support him during his assessment of the current traffic situation (see Sect. 4).

1.1 Related Work

There is a vast set of computer vision algorithms available for the targeted application. The following section gives a brief overview and selects the most appropriate algorithms for the posed challenges.

Lane Recognition

Road coarse estimation is the foundation for the proposed traffic scene analysis on highways. The forward-looking camera image is processed to detect any lane markings. For this purpose, multiple techniques have been introduced [5–9].

They all commonly suggest to place search windows with a width $w(D)$ and height $h(D)$ in dependence on their estimated distance D along the predicted path of the lane. Afterwards, the fit rate between search window hypothesis and image data is analyzed. To increase the detection rate, the search windows are slanted on the basis of the expected curvature of the lane marking (see Fig. 2).

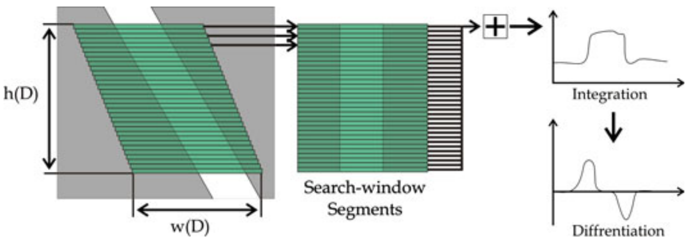


Fig. 2 Lane recognition with slanted and scaled search window segments. The lane is summed up line-wise to generate the integration signal. Afterwards, the differential signal shows the beginning and the end of the road marking

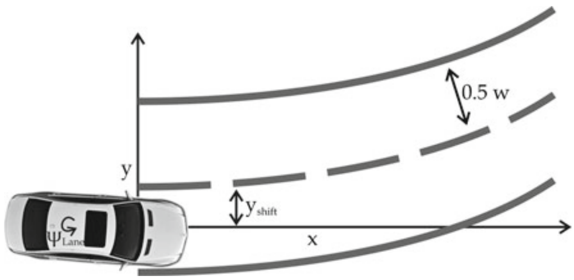


Fig. 3 Vehicle coordinate system and parameters of the lane describing clothides generated by the lane detector

Subsequently, noise is filtered by integrating the grayscale values inside the search window line-wise. To identify the inner and outer boarder of the lane marking, the integrated signal is differentiated and a parabolic curve fit is applied.

By evaluating multiple hypotheses (search window configurations) at different locations with different slants and distances, the best fitting window configuration for the current position is identified.

To transfer the detection results in a lane prediction model, Dickmanns et al. proposed a clothoide curve description [10]. Based on the detection results, the parameters for the lateral shift y_{shift}, the lane width w, vehicle orientation ψ_{lane} as well as the clothoide parameters y_0 and y_1, which describe the curvature, are estimated (see Fig. 3).

$$y_{lane}(x) = \pm 0.5 \cdot w - y_{shift} - \psi_{lane} \cdot x + \frac{1}{2} \cdot y_0 \cdot x^2 + \frac{1}{6} \cdot y_1 \cdot x^3 \qquad (1)$$

The resulting clothoide model encodes road course, vehicle orientation, and relative position. The model is updated iteratively with each timestep t. Naturally, the accuracy of the estimated lane marking position increases from the most distant points towards the position directly next to the vehicle. In consequence, only the latter will be utilized to perform the back-propagation of the road course into the rear-view camera image.

Vehicle Detection and Verification

The road course estimation already yields a static representation of the current vehicle surrounding on a highway. To populate this static model with the dynamic objects, a vehicle *detection* is proposed. Due to its importance for autonomous driving functions, vehicle detection has been a topic of intense research for several years [11, 12].

Initial proposals for identification methods to detect vehicle structures in images were based on hand-crafted features such as contours [13], symmetry [14], corners [15], or edges [16]. However, these algorithms never achieved the detection rates of automatically trained detectors which use a sliding window scheme for hypothesis generation [17–20]. Most of the learned sliding window schemes use cascades of increasingly complex *AdaBoost* classifiers [21, 22] to speed up the detection.

To further increase detector performance, several techniques to optimize the exhaustive sliding window search with improved search patterns have been proposed. These utilize pre-knowledge of the scene [23] or local classifier responses [24].

The applied detector for the results generated is adopted from Gabb et al. [12]. It combines several optimizations: It confines the search to vehicles on an assumed ground plane, uses a boosted cascade with inter-stage information transfer, distance dependent feature sampling and an advanced classifier architecture.

Initially, the hypotheses for the cascaded detection have to be generated. A trivial approach is an exhaustive search based on sliding windows. Following the proposed optimization stated in the previous paragraph, the detector by Gabb et al. reduces the number of hypotheses by assuming a flat ground plane.

Although the number of search windows is reduced, there is still a vast amount of hypotheses to test. Thus, the proposed detector implements a *Boosted Cascade* [21] using *Haar* [25] and *Edge Oriented Histograms* [26] as features to reject non-vehicle hypothesis as early as possible in the detection cascade. To further speed up the search, *Activation History Features* [27] is used to transfer information between the cascade stages.

To reduce the number of hypothesis even further, a multiscale coarse-to-fine search is proposed by Gabb et al. Therefore, the ground plane assumption is extended by taking the effect of foreshortening (see Sect. 7.2 of chapter "Human Visual Perception") into account. Accordingly, the vehicle frontal structures are expected to shrink proportional to their distance and therefore proportional to the vehicle position in relation to the image height. This property is utilized by an image height dependent re-sampling of the detector input. Lower image regions are downsampled whereas image regions showing far objects are processed at full resolution. This approach reduces the number of processed pixels by ≈50 % and yields a speed-up in detector performance by roughly the same factor.

Usually, the detector output confirms multiple nearby hypothesis windows for a single vehicle. As a consequence, in a standard traffic scene there are still a few hundred windows which have to be merged to a single verified detection per vehicle. The state-of-the art algorithms for this post-processing step of vehicle *verification* are reviewed in [23].

As proposed by Gabb et al. the applied classification is based on *Local Receptive Fields* with feed-forward *Neural Networks* which have been proven to have a good performance [28]. They bring the benefit of a more precise object localization. However, the bounding box enclosing the vehicle is still temporal unstable with respect to its location and size in consecutive video frames.

Thus, the vehicle position and contour have to be refined even further. The following section introduces segmentation techniques which are proposed to gain a stable track and a precise vehicle contour, which is required for the targeted application.

Object Segmentation

Image segmentation is the process of partitioning an image into two or multiple segments of pixel groups. This technique is applied to separate detected vehicles from the image background. The known algorithms for this task can be divided into two categories.

The majority of segmenting algorithms rely on contrasts and local contour information at the boundaries of the object to be segmented [29, 30]. Thus, for these algorithms, the information in the segment is disregarded.

The second category are region based techniques which found their segmentation on the information in the object region [31]. Usually, this is implemented by growing a region from "seeds" by iteratively adding similar neighboring pixels to the segment. These algorithms usually suffer from "leaking" in which their greediness leads to a crossing of the object contour.

The *Graph Cut* algorithm [32] combines both properties by applying a global graph cut based cost reduction method. Therefore, a graph with two terminals is build up. The "object" terminal (source) and the "background" terminal (sink) are connected to the marked foreground or background nodes, respectively. These marks have to be generated previously by user interaction (e.g. inpainting). Additionally, the pixels (nodes) are interconnected among themselves with weighted joints.

The weights are derived based on the similarity between the pixels (e.g. intensity, gradient, or color). The weight between object and background terminals and the previously defined foreground and background pixels are set to infinity.

The actual segmentation is performed subsequently by determining a cut that separates the source from the sink with the minimum accumulated edge weight (cost) (see Fig. 4).

Rother et al. [33] extended the *Graph Cuts* to the iterative color sensitive *Grab Cut* algorithm. One of the main benefits of this approach is its simplified user interaction. In contrast to the *Graph Cub* which requires to mark multiple foreground and background seeds in the input image, for the *Grab Cut* it is sufficient to

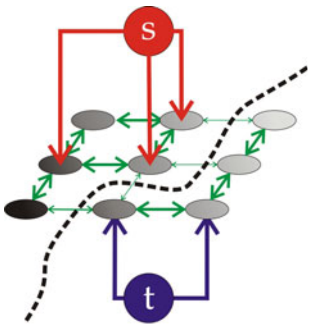


Fig. 4 Image segmentation via *Graph Cut*. The inter pixel weight is encoded in the arrow thickness (*green*). The nodes initially marked as foreground or background are connected with infinite weight to the source (s) or sink (t). The segmentation is given by the mincut of the graph (*dotted line*)

draw a singular rectangle which encloses the foreground object. Thus, the information supplied by the user only describes a clearly defined background and an uncertain region covering foreground and background pixels.

Therefore, the edge weights between the object terminals are no longer infinite since no clearly defined foreground pixels are available from the user input. For this approach, these weights are derived from a color similarity criterion. It generates the initial color signature from the center of the rectangular region and starts an iterative process in which the colors of the object pixels are used to update the signature. Subsequent, a new mincut is performed which results in additional pixels to be identified as foreground. In the next iteration, those pixels are added to the input for the computation of the color signature and the segmentation is repeated.

This iterative mincut algorithm guarantees a proper convergence to at least a local energy minimum for the graph segmentation. The main benefit of this approach lies in the simplified user interaction which does no longer require a precise inpainting of foreground and background segments to achieve a similar quality of segmentation results as *Graph Cuts*.

Hence, the input from the vehicle detection algorithm can be used to automatically segment vehicles from the image background by utilizing *Graph Cuts*. Such a method for automatic vehicle detection, segmentation and tracking is proposed in the following sections.

1.2 Road Course Reconstruction

The road course is reconstructed via an extrinsically and intrinsically calibrated forward-looking camera (see Sect. 1.1). One could argue that the road course can be reconstructed with the backward-looking camera as well. However, a forward-looking orientation enables the algorithms to reconstruct the road course in advance, to build up a stable representation which has its highest accuracy at the vehicles front axis. Due to the ego-motion based translation performed in each timestep t, the highly reliable lane information is propagated behind the vehicle. Beneficially, this translated lane data is not negatively affected by occlusions due to trailing vehicles.

Environment reconstruction is performed by adding the detected lane coordinates next to the front axis $y_{lane}(0)$ to the environment model M in each update step t. An odometry filter estimates the vehicle movement in x direction (Δx) and the rotation around the z axis (ψ_v) for the next time step ($t + 1$). Accordingly, the environment model M has to be translated in x (T_x) and rotated in z (R_z) inverse to the ego-motion of the vehicle (see Eq. 3). This approach ensures that the environment model is kept consistent with the relative position of the vehicle for the next time step $M(t + 1)$.

By applying the projection matrix P to the environment model M, the lane data is projected into the rear-view camera image. The matrix P is built up of the extrinsic (3D translation ($\vec{v}$) and roll (ϕ_c), yaw (ψ_c), pitch (θ_c), angles), and intrinsic (center of distortion u_0, v_0 and focal length f_0) camera calibration data (see Eq. 4).

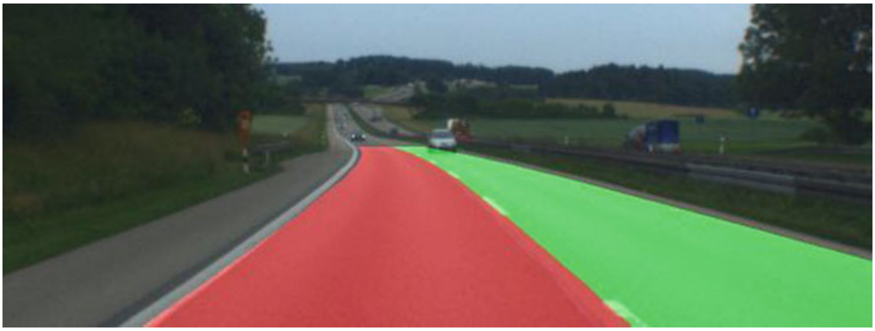


Fig. 5 Rear-view camera image with re-projected road course detected by the forward-looking camera

To visually inspect the coherence, the single re-projected roadway markings may be connected to a closed surface and overlaid to the rear-view camera image (see Fig. 5). This depiction is utilized as well as basis for the time-to-impact visualization presented in Sect. 4.2.

$$M(t) = \begin{pmatrix} x_0 & x_1 & x_2 & \dots & 0 \\ y_0 & y_1 & y_2 & \dots & y_{lane}(0) \\ z_0 & z_1 & z_2 & \dots & 0 \\ 1 & 1 & 1 & \dots & 1 \end{pmatrix} \tag{2}$$

$$M(t+1) = M(t) \cdot T_x(-\Delta x(t+1)) \cdot R_z(-\psi_v(t+1)) \tag{3}$$

$$P = \underbrace{\begin{bmatrix} f_0 & 0 & u_0 & 0 \\ 0 & f_0 & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}}_{intrinsic} \cdot \underbrace{\underbrace{R_z(-\psi_c)}_{yaw} \cdot \underbrace{R_y(-\theta_c)}_{pitch} \cdot \underbrace{R_x(-\phi_c)}_{roll} \cdot T(-\vec{v})}_{extrinsic} \tag{4}$$

1.3 Vehicle Detection and Segmentation

The visualization of the road course in the rear-view camera image is able to support the driver to assess the relative position of the trailing vehicles. Based on this information, some rough estimates of the spatial scene arrangement can be already performed (e.g. road course trajectory, relative position of the own vehicle to the lanes, etc.). However, without any information on position and velocity of the following vehicles, the most important data to assess the current traffic situation is still missing.

The presented framework tackles this problem by extracting the relevant information from the rear-view camera video stream. Therefore, a robust classifier

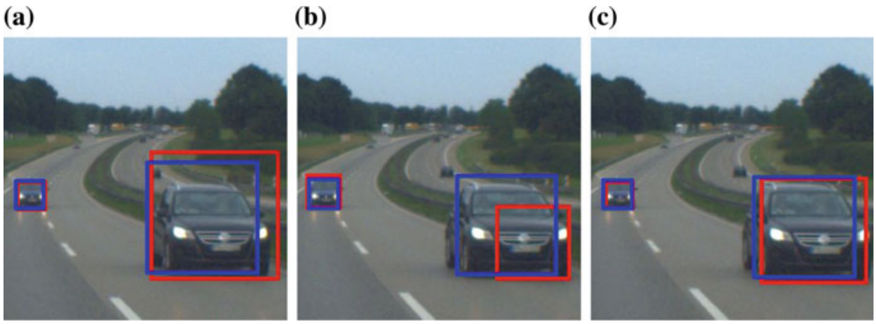


Fig. 6 Unsmoothened vehicle detections (*red*) and temporal smoothed bounding boxes (*blue*). The temporal smoothed boxes are more temporal stable and accurate. **a** Frame 1. **b** Frame 2. **c** Frame 3

based vehicle detection algorithm as presented in Sect. 1.1 is utilized. To generate precise contours of the detected vehicles and to refine the detection results, the *Grab Cut* [33] segmentation algorithm is applied to the bounding box output of the detector (see Sect. 1.1). Besides vehicle position refinement, the precise contour is the basis for the proposed perceptual motivated depictions of the vehicle motion and distance (see Sect. 4).

Vehicle Detection

As first step, the rear-view camera video stream is fed into the proposed vehicle detector by Gabb et al. [12]. To suppress incorrect detections besides the road and to speed up the classifier, the processed image region is restricted to the re-projected road course. The detector output roughly frames each frontal structure of the vehicle with a single bounding box or ROI (Region Of Interest). Unfortunately, as stated in Sect. 1.1, the ROI is temporally unstable with respect to location and size (see red boxes in Fig. 6).

To stabilize the temporal progression of the vehicle detections, the bounding box locations and sizes are smoothed throughout time. This is implemented by associating each subsequent detection at nearly the same location (25 % overlap) with the previous detection from the last frame. Afterwards, an exponential moving average [34] of the associated boxes in width w, height h, and location u, v with a smoothing factor $\delta = 0.3$ is performed:

$$\vec{b}_{det} = [u\, v\, w\, h]^T \tag{5}$$

$$\vec{b}'_{det}(t) = \delta \cdot \vec{b}_{det}(t) + (1 - \delta) \cdot \vec{b}_{det}(t - 1) \tag{6}$$

As a result, the association and temporal smoothing of the detections has the additional benefit that a temporal track of vehicle detentions is built up (see blue boxes in Fig. 6). This 2D track is utilized for the linear motion estimate which is required for the interpolation of single missing detections, and for plausibility checks of the segmentation results (see next section).

Vehicle Segmentation

To perform a natural rendering of supportive visual effects, the precise outline of the detected vehicles has to be estimated. This outline is computed with a segmentation algorithm which separates the vehicle pixels from the background. Additionally, the segmentation is used to achieve a more precise tracking and to determine the exact groundmark of the vehicle on the road surface.

The *Grab Cut* algorithm introduced in Sect. 1.1 requires a coarse mask enclosing the entire object to be segmented. This mask defines a predefined background area and an uncertain region containing both foreground and background pixels (bimap). The *Grab Cut* algorithm shrinks this initial segmentation box to a precise contour of the vehicle. Therefore, it is crucial that the initial bounding box encloses the whole vehicle structure. However, the ROI generated by the detector usually covers only a part of the entire vehicle (see Fig. 6). To compensate for this, the ROI has to be enlarged first. The 2D projection of a vehicle in the rear-view camera resembles a trapezoid shape. Thus, the bounding box is morphed accordingly and its size is increased by 54 % (see red trapezoid in Fig. 7). Finally, the new trapezoid shaped box is used as input for the *Grab Cut* algorithm, which performs a separate segmentation for each vehicle (see Fig. 7).

To gain precise vehicle contours for the subsequent rendering, the segmentation has to be repeated for each new frame. This iterative execution of the segmenting algorithm is utilized to interpolate missing detections. By exploiting the high sample rate of our setup (30 frames per second) the vehicle position does not shift dramatically even if ego- and vehicle-motion is taken into account. Accordingly, the new bimap is generated out of the segment from the previous frame. This is performed by enlarging the previous segment about 30 %. Afterwards, segmentation is repeated to obtain the new contour of the tracked vehicle (see Fig. 7).

In some cases, the algorithm fails either by segmenting the whole masked region as foreground, or by segmenting only a small part of the vehicle. These failure cases can be easily detected by comparing the covered area of the initial bimap and the

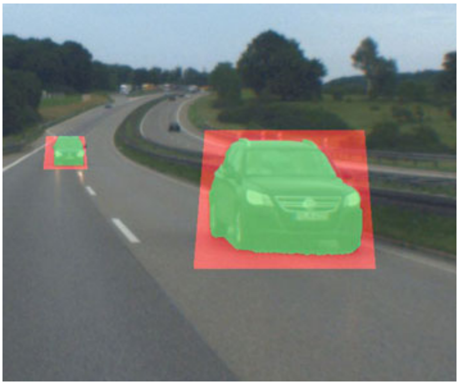


Fig. 7 *Grab Cut* input bimap derived from detector results (*red*) and segmentation result (*green*)

segmentation result. A coverage below 60 % or above 90 % is considered to be invalid. In this case, a second segmentation pass with a new bimap is executed.

The new map is generated out of the vehicle detection result. If there is no detection available, the enlarged bounding box result from the previous frame is used as input bimap. If the second pass fails as well, the previous outline is propagated out of the 2D track generated during the detection. The new position is interpolating based on the averaged 2D shift $\Delta\vec{b}'_{det}(t)$ extracted from the track in image space.

Since the tracking is performed based on iterative *Grab Cut* segmentation of the same vehicle in subsequent frames, it has to be verified that these segments still contain the target vehicle. This validation is performed by associating detections to segments. In case a sequence of segments is not associated with a detection within 10 consecutive frames, the segment is deleted. In case a detection cannot be associated with an existing segmentation, a new track is initialized.

Thus, the presented *Grab Cut* boosted detection method is simple but very effective to track vehicles in video streams. Furthermore, the precise contour segmentation is an essential prerequisite for the plausible rendering of motion and distance cues (see Sect. 4). Additionally, a precise vehicle ground mark is provided, which is utilized for an improved distance and velocity estimate.

1.4 Distance and Velocity Estimation

With the precise vehicle contour available, the ground mark can be assumed to be the undermost pixel of the mask. Based on the flat world assumption ($z = 0$) and the intrinsic and extrinsic camera calibration (see Sect. 1.2), the 3D coordinates (x_g, y_g, z_g) of the ground mark in image space (u_g, v_g) are calculated:

$$\begin{pmatrix} x_g \\ y_g \\ z_g \end{pmatrix} = \begin{pmatrix} \frac{z_0}{\tan\left(\arctan\left(\frac{v_g - v_0}{f_0}\right) + \theta\right)} + x_0 \\ \frac{u_g - u_0}{f_0} \cdot (x_g - x_0) + y_0 \\ 0 \end{pmatrix} \tag{7}$$

The reconstructed distance of the vehicle in combination with the 2D track in image space and the sampling interval Δt_i are utilized to estimate relative speed $\hat{v}_t$ and time-to-impact ΔT_i of the trailing vehicles. To smooth this estimate, an alpha-beta-filtering of consecutive ground mark locations is performed ($\alpha = 0.5$, $\beta = 0.1$) (see Eqs. 8 to 11).

$$\hat{x}_t = \hat{x}_{t-1} + \Delta t_i \cdot \hat{v}_{t-1} \tag{8}$$

$$\hat{x}_t = \hat{x}_t + \alpha \cdot \hat{r}_t \qquad \| \hat{r}_t = x_t - \hat{x}_t \tag{9}$$

$$\hat{v}_t = \hat{v}_{t-1} + \frac{\beta}{\Delta t_i} \cdot \hat{r}_t \tag{10}$$

$$\Delta T_i = \frac{\hat{x}_t}{\hat{v}_t} \tag{11}$$

As last step, vehicle tracks and velocity as well as the distance estimate are merged with the current environment model M. This meta-data is attached to each frame and fed into the subsequent rendering pipeline (see Sect. 4). This combined data does not only provide the spatial and temporal reconstruction of the rear-ward traffic scene but with the time-to-impact the risk potential of a collision for each trailing vehicle is given as well.

1.5 Monocular Scene Reconstruction Results and Evaluation

The accuracy of the provided risk potential assessment and depth reconstruction relies on the precision of the estimated vehicle distance and road course estimate. Therefore, it is proposed to compare the estimated depth with the distances measured by a stereo camera system. To allow a direct comparison, a similar secondary camera was mounted besides the monocular backward-looking camera. The resulting reference stereo system had a base-width of 25 cm. The left camera of this system was fed into the monocular processing pipeline whereas the stereo pair has been processed with the presented *Semi Global Matching* algorithm [2].

For the monocular toolchain, depth maps have been estimated based on the reconstructed environment model M. The depth for pixels marked as road have been computed in accordance to formula 7. For pixels that were part of a segmented vehicle, the distance value of the ground mark was assigned. All other pixels were set to infinite distance (see Fig. 8). However, this approach implicitly adds the erroneous assumption of an invariant depth for the visible vehicle structures. Although this results in slightly wrong estimates, the generated depth map is suitable to render plausible visual effects (see Fig. 14).

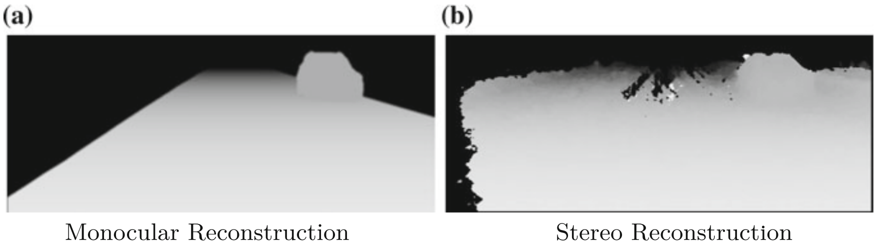


Fig. 8 **a** Depth map reconstructed from mono camera based environment reconstruction, **b** depth map computed with Semi Global Matching

The proposed depth reconstruction methods processed 60 frames from three different highway sequences. For both approaches, the left camera was chosen as origin of the coordinate frame for the depth calculations.

As error measure the absolute mean distance between the reconstructed disparity map generated by the presented approach and the disparity map from the stereo system was calculated. If one of the two evaluated methods had no valid data at the compared position, the values were discarded. The resulting mean absolute error across all evaluated scenes was $\bar{x}_{err} = 1.95\,\mathrm{m}$.

The incorrect depth estimates have various sources. The largest errors in monocular reconstruction are induced by pitching and rolling of the vehicle. Additionally, uneven surfaces violate the flat world assumption and induce further mismatches. Smaller but perceivable errors are induced by the simplification of plane vehicle structures (see Fig. 8).

These errors and the missing distance information for unknown objects (e.g. objects next to the road or unclassified objects) disqualify this technique for most computer vision applications supporting automatic driving functions. Nonetheless, the obtained distance and edge accuracy of the depth map are sufficient to produce high quality renderings for intuitive distance, velocity and risk potential visualizations (see Sect. 4).

Due to the fact that many modern upper-class vehicles are already equipped with forward and backward-looking cameras, the implementation of the proposed traffic analysis is a logical next step. The resulting system enables to assess risk potential, time-to-impact, relative position with respect to the track, and velocity of all trailing vehicles. This cumulative information on the traffic situation is a rich basis for multiple rendering techniques which are able convey these critical information to the driver in a natural but distinctive way (see Sects. 2 and 4).

2 Artificial Depth-of-Field Rendering

Most small consumer (e.g. for cell phones) as well as automotive cameras are built up of small image sensors and compact lenses with a very small aperture and short focal length. Images captured with such cameras usually do not show any defocus blur for larger distances than several centimeters to meters. From this it follows that these images render a nearly uniform sharpness within the entire image.

Thus, with a depth map available, it is feasible to render an artificial depth-of-field into these overall sharp input images.

2.1 Mathematical Description of Approximated Defocus Blur

To introduce a plausible depth-of-field into a sharp input image, the underlying optical effects of defocus blur have to be considered. An ideal thin lens focuses

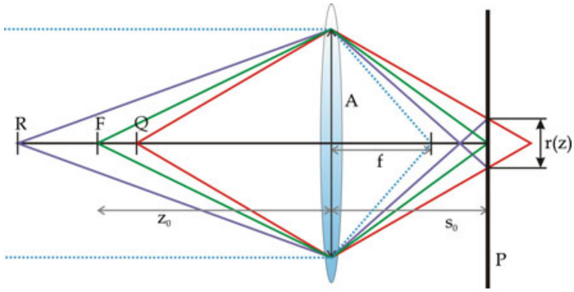


Fig. 9 Lens Projection: Objects at distance z_0 are projected sharply onto the imager plane P. Objects at other distances are projected as blurred *circles* with a diameter size of $r(z)$

incoming parallel rays to a single point on the opposite side of the optics. The distance between lens and the point of collineation is called focal length f. If the distance between image plane and lens equals s_0, light emanating from the focal distance z_0 will focus on the image plane. Objects at other distances will not collineate at a single point. Therefore, light emanating from out of focus distances z_x are widened to a spot, which is defined as *Blur Circle* $r(z)$ (see Fig. 9).

$$r(z_1) = \left| A \frac{s_0}{z_0} \left(1 - \frac{z_0}{z_1} \right) \right| \tag{12}$$

The increase in the blur circle diameter is gradual but not symmetrical for farther and closer distances than z_0 (see Fig. 10). The lower threshold for the distance rendered with acceptable sharpness shall be z_l whereas the outer boarder shall be z_u. The depth-of-field describes this blur spectrum between z_l and z_u in which the image still appears to be sharp due to the limited optical resolution of the eye and/or displaying device. Thus, the upper limit is not a fix value but is related to the resolution of the capturing and displacing device as well as the viewing distance and the human visual acuity (see Sect. 5 of chapter "Human Visual Perception").

Due to the same physiological limitations, deviation in sharpness of objects above a distance of z_∞ is not perceivable as well. From this it follows that if

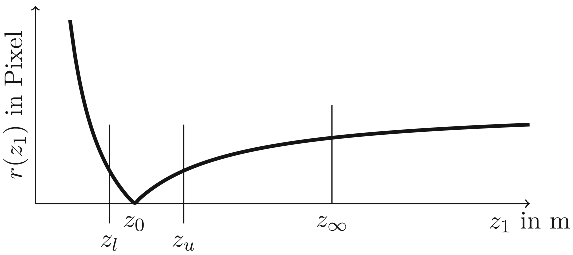


Fig. 10 Blur radius $r(z1)$ in dependency of object distance z_1 with the focus distance z_0. At distances above z_∞ the image appears to have a uniform sharpness

$z_u >= z_\infty$, the deviations in sharpness for distances above z_l cannot be perceived. This effect can be used to tune an optical system to show a uniform sharpness for distances above z_l.

To give a numerical example, a typical automotive camera has an imager width of ≈ 0.85 cm with a lens aperture A of $f/2$ at a focal length f of 4 mm focused at 6 m. This setup generates from 1.5 m up to infinity a nearly uniform sharpness. On the upside, this removes the requirement of any mechanical parts for refocusing inside the camera. On the downside, the very supportive monocular depth cue of blur-from-defocus as described in Sect. 7.2 of chapter "Human Visual Perception" is not present.

However, with the underlying depth map available, the required blur radii $r(z1)$ for any focus distance z_0 can be calculated and used to emulate lenses with various apertures.

To emulate a natural depth-of-field, that supports the viewer in his assessment of inter scene distances, the optical parameters of a human eye have to be

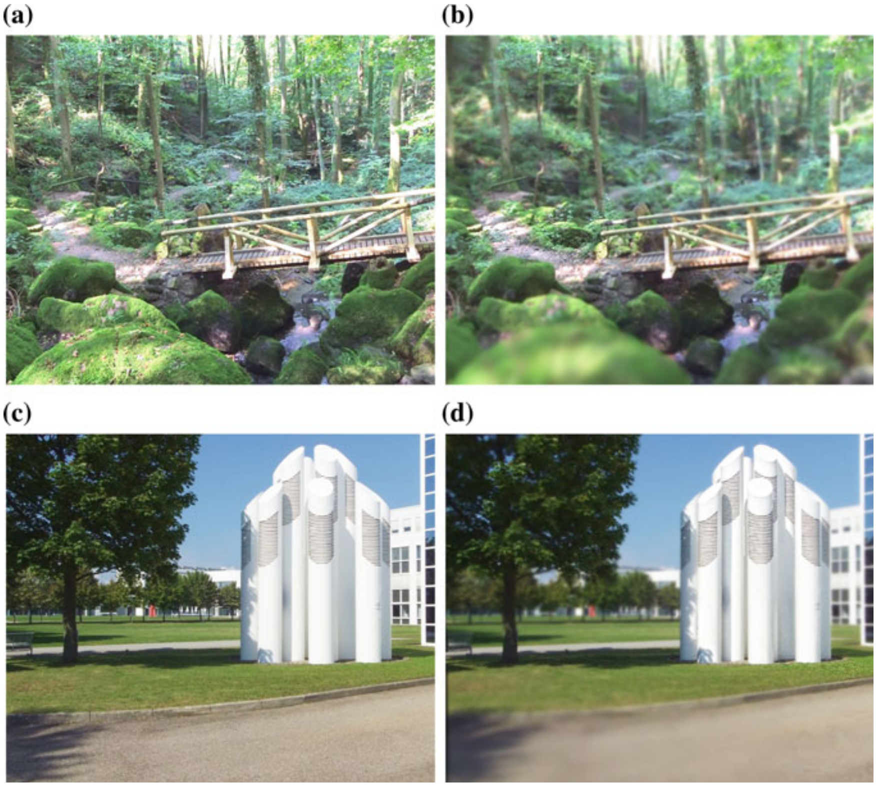


Fig. 11 Real-Time depth-of-field renderings performed with à -trous wavelets. **a** Original Bridge. **b** Depth-of-field rendering. **c** Original Vent. **d** Depth-of-field rendering

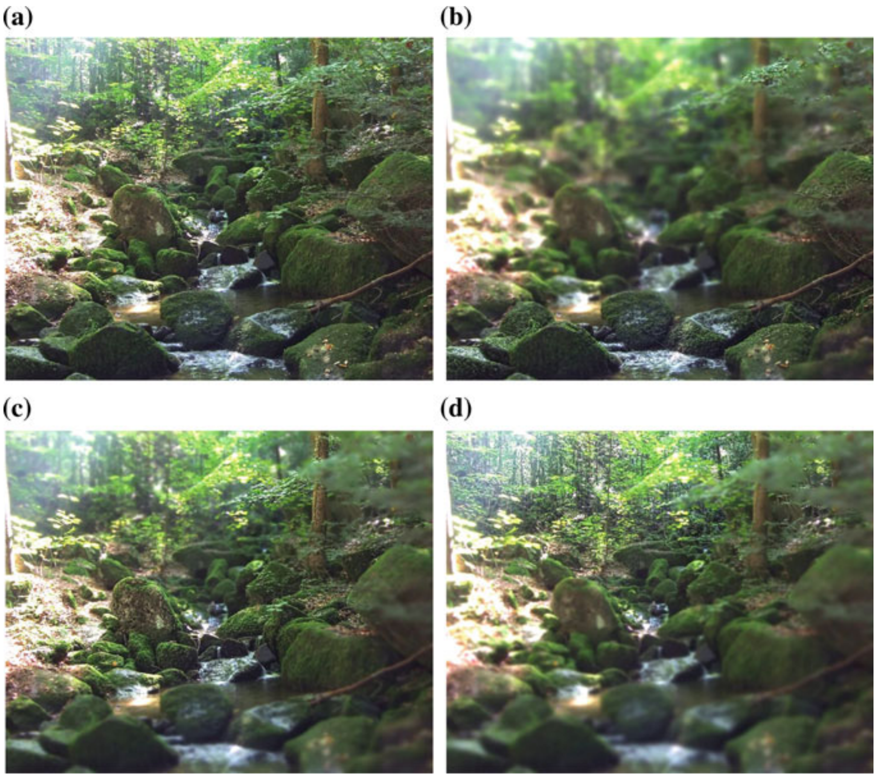


Fig. 12 Three different focal distances emulated with the proposed artificial depth-of-field rendering. Shot was taken with a *Fuji W1* consumer stereo camera. **a** Original. **b** Near Focus. **c** Medium Focus. **d** Far Focus

replicated. Following the properties described in Sect. 2 of chapter "Human Visual Perception" the average eye aperture is $A \approx 4.6\,\text{mm}$ with a focal distance of $s_0 \approx 24\,\text{mm}$.

To render an approximation of a physically correct defocus blur, each pixel in the input image has to be convoluted with a Gaussian kernel of the size $r(z_1)$. For example, to emulate the human eye depth-of-field for a camera with a pixel size of 4.2 µm for $z_0 = 0.5$ m and $z_1 = 10$ m requires blur kernel sizes between 0 and 25 pixels. Fast techniques to perform DOF renderings can be found in [35].

Resulting example images with blur-from-defocus depictions are presented in Figs. 11 and 12.

The resulting images have a more spacial and pleasant appearance for viewers. The latter is underpinned by the study by Hillaire et al. [36]. Additionally, findings in perceptual science indicate that the viewer's attention is drawn by high frequency edges of any focused object, which may be utilized for user attention guidance.

3 Distance-Adaptive Desaturation

Besides the previously proposed depth-of-field renderings, manipulations of color can be utilized to induce depth sensation as well.

The human visual system extracts depth information from various physical effects which manifest as pictorial depth cues. One of these effects is atmospheric haze which has been discussed in Sect. 7.2 of chapter "Human Visual Perception". Furthermore, the findings of Troscianko et al. [37] indicate that depth encoded in color (de-)saturation influences human depth perception in general. These results point towards a subconscious interpretation of relative color saturation in the viewed image. This effect is associated with the weak ordinal depth cues which support the assessment of scene layering and foreground/background separation process (see Sect. 5.2 of chapter "Human Visual Perception").

4 Motion and Distance Visualization for Rear-View Camera Applications

The trajectory and relative velocity of trailing vehicles is very essential information to assess the traffic situation on a highway. To accommodate for this, a supportive enrichment with motion and distance cues of a rear-view camera video stream is proposed.

A trivial implementation to convey this information to the driver could be a simple overlay which shows speed, distance, and fast approaching vehicles as textural information and highlighting frames (see Fig. 13). However, such visualizations lead to numerous overlayed data which might overextend the viewer. Especially while driving on a highway, an intuitive visualization of the critical information is preferable.

Fig. 13 Rear-view camera image with distance and velocity overlays. The information display may confuse the driver

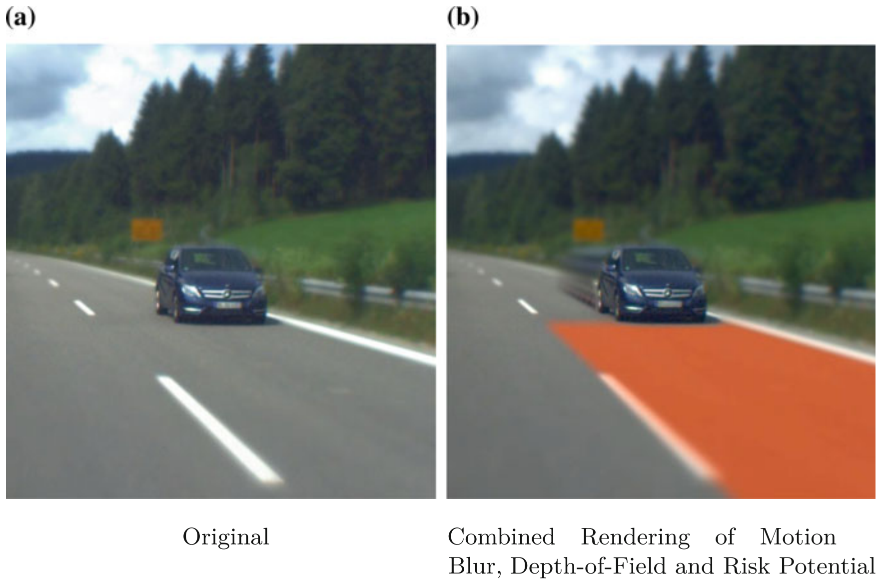


Fig. 14 **a** Original. **b** Shows a combination of motion blur, depth-of-field and time-to-impact visualization. Therefore, within a single image supplemental information on approaching speed, distance and time-to-impact is conveyed

Videos captured with a vehicle mounted camera usually show large displacements within the entire scene. However the optical flow originating from the camera's ego-motion has to be separated from the flow induced by the motion of the trailing vehicles. Hence this is a computational task, which is know to be challenging [38].

A more efficient but specialized method for the traffic scene reconstruction was presented in Sect. 1. Although it is confined to highway scenarios and implies a flat world, the edge coherence for the segmented vehicles is high and its distance accuracy is adequate for visualization applications.

Three depictions to support the driver's assessment of the scene are proposed. To convey the distance information, the previously presented *artificial depth-of-field rendering* is applied. Additionally, to emphasize a high approaching speed of trailing vehicles, *exaggerated artificial motion blur* is rendered. Furthermore, the new visualization of *color coded risk potential* is proposed. The latter depiction visualizes the risk potential of an imminent collision with an approaching vehicle by color-coding the lane in relation to the time-to-impact (see Fig. 14b).

The approaches for the rendering of an artificial depth-of-field are evaluated in Sect. 2. The implementation of additional visualizations is presented in the following two sections.

4.1 Exaggerated Artificial Motion Blur Rendering

The rendering method for introducing motion blur into image material is adopted from Potmesil and Chakravarty [39]. Fast approximations for these kind of renderings have been presented in [1].

To execute the algorithm, the following steps are taken. Initially, the apparent shift between two frames $\vec{m}_d(t)$ of the detection box center coordinates (u_d, v_d) is calculated (see Eq. 13). To create an exaggerated motion blur, the flow vector $\vec{m}_b(t)$ is prolonged by the factor γ which was set to $\gamma = 2.0$ for all presented renderings.

Subsequently, the motion vector is assigned to all pixels $(p(x_0, y_0))$ which are elements of the segmented vehicle (see Eq. 15). Afterwards, the segmented pixels are copied into a new image q and translated along the motion trajectory. The parameter K controls for the granularity of the rendering (for the presented results $K = 15$). The intensity of the multiple copies is faded linearly with growing distance $(\frac{k}{K})$ (see Eq. 16).

Finally, the multiple copies of the moving vehicle are convoluted with a discrete Gaussian kernel along the motion trajectory $(q * g)$ with $\sigma = 1$. In the output image the motion trail rendered in q is blended on p whereby pixel positions segmented as vehicle are excluded (see Figs. 14 and 15). The resulting rendering introduces an exaggerated motion blur of the trailing vehicles to depict their approaching speed in an intuitive manner. Unfortunately, the utilized implementation has a long runtime of 1.44 s for the rendering of 640 × 480 pixel in a CPU implementation.

$$\vec{m}_d(t) = \begin{pmatrix} u_d(t) \\ v_d(t) \end{pmatrix} - \begin{pmatrix} u_d(t-1) \\ v_d(t-1) \end{pmatrix} \tag{13}$$

$$\vec{m}'_d(t) = \gamma \cdot \vec{m}_{dd}(t) \tag{14}$$

$$\vec{m}(x_0, y_0) = \vec{m}'_d(t) \quad | p(x_0, y_0) \in \text{segment} \tag{15}$$

Fig. 15 Artificial exaggerated motion blur rendering to support the driver's assessment of the approaching speed and trajectory of the trailing vehicle

$$q\left(x_0 - m'_x\left(1 - \frac{k}{K}\right), y_0 - m'_y\left(1 - \frac{k}{K}\right)\right) = p(x_0, y_0) \cdot \left(\frac{k}{K}\right) \quad |\vec{m}'(t, x_0, y_0) \neq 0 \quad (16)$$

$$g(m) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{m^2}{2\sigma^2}} \quad (17)$$

$$q'(x_0) = \sum_{m=-M}^{M} g(x_0 - m) \cdot g(m) \quad (18)$$

4.2 Risk Potential Visualization

The presented depictions of depth (see Sect. 2) and velocity enrich the output image with subconscious cues to convey critical information on the current traffic situation to the viewer. However, the risk of a collision during lane change shall be illustrated as well. Due to the urgency of this information, the visualization has to be prominent and yet intuitively understandable.

The findings of Karowski et al. [40] indicate that a large amount of red colors in an image induces higher neuronal activity and therefore the strongest response for human perception (see Sect. 3.1 of chapter "Human Visual Perception"). Their studies show that level of arousal gradually decreases from red with the colors orange through green to white. Based on these findings, a hazard-potential controlled color-coding is introduced (Fig. 16).

The proposed algorithm blends signal colors onto the lane of the approaching vehicle to intuitively convey the time-to-impact ΔT_i. The rendering is implemented by determining distance and track of the trailing vehicle. Afterwards, the associated lane overlay is calculated as described in [1] and shortened to the distance of the vehicle's groundmark. The color of the blended overlay is set as proposed by Karowski et al. to colors which are gradually faded from red through orange to green (see Fig. 16). As a result, the manipulated image supports the driver to assess the risk potential of a collision while changing lanes on a highway.

(a) **(b)**

Fig. 16 Time-to-impact dependent color-encoded lane overlay to intuitively visualize the risk potential of a lane change to the driver

4.3 Perceptual Study on Exaggerated Motion Blur

Asides of the motion blur rendering techniques, the impact of the generated images on the HVS (Human Visual System) has to be evaluated. Burr and Ross [41] have found evidence that exaggerated motion blur (speed-lines) has a direct influence on human velocity assessment. Additionally, the perceptual findings reviewed in Sect. 4 of chapter "Human Visual Perception" as well as the physiological properties of the photoreceptor cells (see Sect. 4.1 of chapter "Human Visual Perception") support the hypothesis that human vision subconsciously processes motion blur. To evaluate if these findings apply to the proposed enrichment of rear-view camera streams, a comparative study was performed.

The main goal was to assess the feasibility of inducing the sensation of a high relative vehicle velocity by introducing exaggerated motion blur. Conclusively, the hypothesis was proposed that vehicles rendered with artificial motion blur are overestimated in their velocity—independent from their actual relative speed. Therefore, the study evaluated the velocity assessment for video sequences with artificial motion blur and color encoded risk potential renderings in comparison to the original sequence. The latter rendering was added to examine the correlation between color-coding visualization and the viewer's velocity assessment.

Study Design

Short video sequences (1 s) captured with a vehicle mounted rear-view camera were shown to the participants. The carrying vehicle was driving at constant speed on the right highway lane. To avoid any interference with color or brand of the vehicle, all sequences showed the same trailing car (blue Mercedes-Benz B-class). The short duration of the video sequence limited the influence of conscious processing by the user. Each scenery was only presented once to each participant to eliminate the influence of previous knowledge or learning effects.

After each sequence the user had to decide if the trailing vehicle's relative velocity was lower, the same, higher, or much higher. The users were allowed to skip the question if they were uncertain regarding their answer.

Study Setup

As stimulus data, 52 scenes (8 slower ($\approx -20\,\text{km/h}$), 17 of the same speed ($\approx \pm 0\,\text{km/h}$), 25 faster ($\approx +20\,\text{km/h}$), 2 much faster ($\approx +40\,\text{km/h}$)) were captured. All scenes, including those showing a decelerating car, were rendered with the color-coded time-to-impact, or with the artificial motion blur depiction, respectively. The latter had the main goal to induce a strong sensation of high approaching speed independent from the actual occurrence. The resulting renderings and the original scene were added to a sequence pool (53×3 scenes). Out of this pool 6 or 7 scenes were randomly chosen without repetition and presented to one of the 70 participants.

To familiarize the participants with the system, they were shown 4 different scenes with each of the 4 different velocities of trailing vehicles. The training results were excluded from the evaluation of the study.

Table 1 Chi square analysis of the perceived relative speed of a following vehicle

Visualization	Under-estimated	Correctly estimated	Over-estimated	$\sum_i$
Original (f_e) r^2	24 (23) 0.01	123 (114) 0.70	*9* *(19) 4.90*	**156**
Motion blur (f_e) r^2	17 (23) 1.42	106 (110) 0.17	28 *(18) 5.63*	**151**
Risk potential (f_e) r^2	26 (21) 1.25	97 (102) 0.21	16 (17) 0.01	**139**
$\sum_j$	**67**	**326**	**53**	**446**

The first row of every category indicates the observed frequency (f_o). The second row indicates the expected frequency (f_e) and the quadratic residue r^2. *Blue* colored frequencies show significantly (p = 0.05) lower speed estimation. *Red* colored frequencies show a significantly (p = 0.05) higher estimation. Thus, it is shown that the perceived velocity is manipulated by the motion blur rendering and is therefore suited to convey the information of a fast approaching vehicle

In summary, 477 sequences were shown, 446 answers were valid. Skipped answers (31 in total, 1 original, 26 motion blur, 4 risk potential) were excluded from the evaluation.

Evaluation

A Pearson chi square analysis [42] was performed to evaluate the answers. The postulated null hypothesis was “The perceived approaching speed is not influenced by adding additional motion cues”. The answers were classified in three categories. The perceived relative speed of the following vehicle was underestimated, correctly estimated, or overestimated. The statistic test for significance (see Table 1) is performed by comparing the sum over all residues $\chi^2 = \sum r^2$ with the value for a $\chi^2(\mathrm{df}, p)$ distribution with the same degree of freedom (df = 4) and a high significance ($p = 0.01$) level. The sum $\sum r^2 = 14.34$ exceeds $\chi^2(4, 0.01) = 13.2$ and therefore the null hypothesis has to be rejected. From this one can deduce that the applied visualization has an impact on perception of velocity. For values for which $r^2 > 3.84$ is true the observed frequency differs significantly from the expected frequency. From this it follows that the perceived speed in the original videos was not overestimated by a significant number of participants (9 out of 156) whereas the artificial motion blur visualization led to a significant overestimation of the approaching speed (28 out of 151). The risk potential visualization showed no significance.

From the results one can conclude the following: First, it shall be noted that the participants were quite reliable in the correct assessment of the vehicle’s relative speed. The results show that a significant number of participants did not tend to overestimate velocity. Furthermore, the study shows that a significant number of users overestimated the speed of the vehicles rendered with artificial motion blur (see Fig. 17). From this it follows that artificial motion blur is able to induce an overestimation of the assessed velocity of vehicles in video streams.

Additionally, it is shown that the risk potential rendering has no impact on the assessment of velocity in video sequences. Thus, this rendering has only an alerting character but does not influence the driver’s velocity estimation.

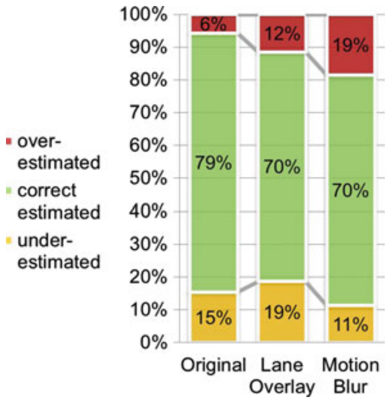


Fig. 17 Relative probands velocity estimation results. The study shows that artificial motion blur induces a statistically significant over-estimation of velocity, whereas the lane overlay does not show a statistical relevant deviation from a normal distribution

5 Conclusion

The previous chapter introduced a novel technique for monocular camera based temporal and spatial scene reconstruction for rear-view camera systems as input for motion and distance visualizations in still and video images.

The depiction and enhancement of monocular depth cues in 2D images enriches the output material with distance information. Thus, the resulting still and video images are able to subconsciously transport depth, scene layering, and saliency information without disrupting the natural appearance of the input image.

Furthermore a secondary rendering method to generate artificial exaggerated motion blur for approaching trailing vehicles in a rear-view camera applications was presented. The related perceptual study revealed that vehicles rendered with artificial motion blur induce the sensation of a high approaching speed—independent from their actual relative velocity. Thus, it is shown that these depictions are able to alter the mental representation of the viewed scene. If such renderings are only applied to fast approaching vehicles, the driver's awareness of the current traffic situation can be increased.

Hence, founded on the perceptual motivated rendering techniques distance, velocity and risk potential may be conveyed to the driver in a natural non-disruptive way. Accordingly these visualizations are able to support the assessment of traffic situations to mitigate the risk of collisions while minimizing driver distraction.

References

1. Rößing, C.: Video and image manipulation for enhanced perception, Dissertation, University Ulm, Online. http://vts.uni-ulm.de/doc.asp?id=9319(2014)
2. Hirschmüller, H., Innocent, P.R., Garibaldi, J.: Real-time correlation-based stereo vision with reduced border errors. Int. J. Comput. Vis. **47**(1): 229–246, issn: 09205691. doi:10.1023/A:1014554110407 (2002). Online http://www.springerlink.com/content/72d01vcdjd798bay/
3. Huang, Y.-W., Chen, C.-Y., Tsai, C.-H., Shen, C.-F., Chen, L.-G.: Survey on block matching motion estimation algorithms and architectures with new results, Journal of VLSI signal processing systems for signal, image and video technology, **42**(3): 297–320, issn: 0922-5773. doi: 10.1007/s11265-006-4190-4 (2006). Online http://dx.doi.org/10.1007/s11265-006-4190-4
4. Nedevschi, S., Schmidt, R., Graf, T., Danescu, R., Frentiu, D., Marita, T., Oniga, F., Pocol, C.: Proceedings of 3d lane detection system based on stereovision. The 7th International IEEE Conference on Intelligent Transportation Systems (IEEE Cat. No.04TH8749) (2004). doi: 10.1109/ITSC.2004.1398890
5. Mysliwetz, B.: Parallelrechner-basierte Bildfolgen- interpretation zur autonomen Fahrzeugsteuerung, PhD thesis, Universität der Bundeswehr München (1990)
6. Gern, A.: Multisensorielle Spurerkennung für Fahrerassistenzsysteme, PhD thesis, Universität Stuttgart (2005)
7. Goldbeck J., Huertgen, B.: Lane detection and tracking by video sensors. In: Proceedings of 199 IEEE/IEEJ/JSAI International Conference on Intelligent Transportation Systems (Cat. No.99TH8383), (1999). doi: 10.1109/ITSC.1999.821030
8. Risack, R., Klausmann, P., Kruger, W., Enkelmann, W.: Robust lane recognition embedded in a real time driver assistance system, IEEE International Conference on Intelligent Vehicles, pp. 35–40 (1998)
9. Franke, U., Loose, H., Knoppel, C.: Lane recognition on country roads, in Intelligent Vehicles Symposium, IEEE, pp. 99–104 (2007). doi: 10.1109/IVS.2007.4290098
10. Dickmanns, E.D., Zapp, A.: A curvature-based scheme for improving road vehicle guidance by computer vision, In: Proceedings of SPIE, vol. 0727, pp. 161–168. doi: 10.1117/12.937795 (1987). Online http://dx.doi.org/10.1117/12.937795
11. Sivaraman S., Trivedi, M : Integrated lane and vehicle detection, localization, and tracking: a synergistic approach, IEEE Trans. Intel. Trans. Syst., **14**(2), 906–917, issn: 1524–9050. doi: 10.1109/TITS.2013.2246835 (2013). Online http://ieeexplore.ieee.org/xpls/abs
12. Gabb, M., Löhlein, O., Wagner, R., Westenberger, A., Fritzsche, M., Dietmayer, K.: High-performance on-road vehicle detection in monocular images. In: Proceedings of ITSC (2013)
13. Handmann, U., Kalinke, T., Tzomakas, C., Werner, M., Seelen, W.: An image processing system for driver assistance, Image Vis. Comput., **18**(5), 367–376, issn: 02628856. doi: 10.1016/S0262-8856(99)00032-3 (2000). Online. http: //linkinghub.elsevier.com/retrieve/pii / S0262885699000323
14. Broggi, A., Cerri, P., Antonello, P.C.: Multi-resolution vehicle detection using artificial vision, In: Proccedings of IEEE Intelligent Vehicles Symposium, pp. 310–314 (2004)
15. Broggi, A., Bertozzi, M., Castelluccio, S.: A real-time oriented system for vehicle detection. J. *Syst.* Architect. (1997). doi: 10.1016/S1383-7621(96)00106-3
16. Sun, Z. S. Z., Miller, R., Bebis, G., DiMeo, D.: A real-time precrash vehicle detection system. In: Proceedings of Sixth IEEE Workshop on Applications of Computer Vision (WACV) (2002). doi: 10.1109/ACV.2002.1182177
17. Caraffi, C., Vojir, T., Trefny, J., Sochman, J., Matas, J.: A system for real-time detection and tracking of vehicles from a single car-mounted camera. In: Proceedings of (ITSC), 2012 15th International IEEE Conference on Intelligent Transportation Systems, IEEE, pp. 975–982 (2012)
18. Yong, X., Zhang, L., Song, Z., Hu, Y., Zheng, L Zhang, J.: Real-time vehicle detection based on haar features and pairwise geometrical histograms. In: Proceedings of IEEE International

Conference on Information and Automation, pp. 390–395 (2011).doi:10.1109/ICINFA.2011.5949023
19. Felzenszwalb, P., Girshick, R., McAllester, D.: Cascade object detection with deformable part models. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), issn: 1063-6919 (2010). doi:10.1109/CVPR.2010.5539906
20. Gabb, M., Lohlein, O., Oberlander, M., Heidemann, G.: Efficient monocular vehicle orientation estimation using a tree-based classifier, In: Proceedings of IEEE Intelligent Vehicles Symposium (IV), pp. 308–313, issn: 1931-0587 (2011). doi: 10.1109/IVS.2011.5940516
21. Viola, P., Jones, M.: Robust real-time object detection, International. J. Comput. Vis. **57**(2), 137–154, issn: 09205691. doi: http://dx.doi.org/10.1023/B:VISI.0000013087.49260.fb (2001). Online http ://scholar.google.com/scholar ? hl = en
22. Schapire, R.E., Singer, Y.: Improved boosting algorithms using confidencerated predictions, In: Machine Learning, pp. 80-91, isbn: 1581130570. http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.33.4002 (1999)
23. Enzweiler, M., Kanter, P., Gavrila, D.: Monocular pedestrian recognition using motion parallax, In: Proceedings of IEEE Intelligent Vehicles Symposium, issn: 1931-0587 (2008). doi: 10.1109/IVS.2008.4621169
24. Lampert, C., Blaschko, M., Hofmann, T.: Beyond sliding windows: object localization by efficient subwindow search, In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (2008)
25. Oren, M., Papageorgiou, C., Sinha, P., Osuna, E., Poggio, T.: Pedestrian detection using wavelet templates, In: Proceedings of *IEEE Computer Society Conference on Computer Vision and Pattern Recognition,* pp. 193–199, isbn: 0818678224 (1997) doi: 10.1109/CVPR.1997.609319
26. Levi K., Weiss Y., Learning object detection from a small number of examples: the importance of good features.In: Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR) vol. 2, issn: 1063-6919 (2004). doi: 10.1109/CVPR.2004.1315144
27. Gressmann, M., Palm, G., Lohlein, O.: Surround view pedestrian detection using heterogeneous classifier cascades. In: Proceedings of 14th International IEEE Conference on Intelligent Transportation Systems (ITSC), pp. 1317–1324, issn: 2153-0009. (2011) doi: 10.1109/ITSC.2011.6082895
28. Gabb, M., Wagner, R., Gressmann, M., Hartmann, O., Lohlein, O., Schweiger, R., Dietmayer, K.: Feature selection for automotive object detection tasks—a study. In: Proceedings of *IEEE Second International Conference on Consumer Electronics—Berlin (ICCE-Berlin)*, IEEE, pp. 209–213, isbn: 978-1-4673-1547-0. (2012). doi: 10.1109/ICCE- Berlin.2012.6336450 Online. http://ieeexplore.ieee.org/xpl/articleDetails.jsp?tp
29. Kass, M., Terzopoulos, D., Witkin, A.: Snakes: active contour models (1988). doi: 10.1007/BF00133570
30. Mortensen, E.N., Barrett, W.A.: Interactive segmentation with intelligent scissors (1998). doi:10.1006/gmip.1998.0480
31. Haralick R.M., Shapiro, L.G.: Computer and Robot Vision, June. Addison-Wesley, 1992, vol. 1, ch. 316-317, pp. 121-122, isbn: 0201108771. Online. http://www.mendeley.com/research/computer-and-robot-vision-1/
32. Boykov Y., Jolly, M.-P.: Interactive graph cuts for optimal boundary region segmentation of objects in n-d images,In: Proceedings Eighth IEEE International Conference on Computer Vision. ICCV , vol. 1, issn: 0277786X. (2001) doi: 10.1109/ICCV.2001.937505
33. Rother, C., Kolmogorov, V., Blake, A.: Grab Cut, ACM Transactions on Graphics. **23**(3), p. 309, issn: 07300301. doi: 10.1145/1015706.1015720 (2004) Online. http://portal.acm.org/citation.cfm?doid=1015706.1015720
34. Mertens P., Raessler, S.: Prognose Rechung. Physicas, vol. 6, isbn: 3790802166 (2004)

35. Rößing, C., Reker, A., Gabb, M., Dietmayer, K., Lensch, H.: Intuitive visualization of vehicle distance, velocity and risk potential in rear-view camera applications, In: Proceedings Intelligent Vehicles Symposium (IV), IEEE, pp. 579–585 (2013)
36. Hillaire, S., Lecuyer, A., Cozot, R., Casiez, G.: Using an eye-tracking system to improve camera motions and depth-of-field blur effects in virtual environments, Virtual Reality Conference, VR '08. IEEE,**28**(6), pp. 47–50, issn: 0272-1716. (2008) doi: 10.1109/VR.2008.4480749
37. Troscianko, T., Montagnon, R., Le Clerc, J., Malbert, E., Chanteau, P.L.: The role of colour as a monocular depth cue., Vision Research, **31**(11), pp. 1923–1929, (1991). Online http://www.ncbi.nlm.nih.gov/pubmed/1771776
38. Klappstein, J., Vaudrey, T., Rabe, C., Wedel, A., Klette, R.: Moving object segmentation using optical flow and depth information. in: Wada, T., Huang, F., Lin, S. (Eds.) Advances in Image and Video Technology, vol. 5414, pp. 611–623. Springer Berlin, Ch. Lecture No, isbn: 9783540929567. doi: 10.1007/978-3-540-92957-4_53. (2009) Online http://dx.doi.org/10.1007/978-3-540-92957-4
39. Potmesil, M., Chakravarty, I.: A lens and aperture camera model for synthetic image generation (1981). doi:10.1145/965161.806818
40. Karwowski, W.: The discipline of ergonomics and human factors, Handbook of human factors and ergonomics, vol. 3 (2006)
41. Burr, D.C., Ross, J., Morrone, M.C: Seeing objects in motion. In: Proceedings of the Royal Society of London. Series B, Containing papers of a Biological *character*. Royal Society (Great Britain), vol. 227, no. 1247, pp. 249–265, 1986, issn: 0962-8452. doi: 10.1098/rspb.1986.0022
42. Pearson, K.: X. on the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling, Philosophical Magazine Series *5*, vol. 50, no. 302, pp. 157–175, 1900. doi: 10.1080/14786440009463897. [Online]. http://www.tandfonline.com/doi/abs/10.1080/14786440009463897

Functional Safety of Camera Monitor Systems

Bernhard Kaiser

Abstract Although the main concern of developers of Camera Monitor Systems (CMS) is usually about meeting the legal requirements and image quality demands, one should not forget that CMS are electronic systems with the potential to endanger vehicle safety, and therefore subject to a Functional Safety process according to the worldwide standard ISO 26262. Functional Safety is the discipline that systematically discusses hazards arising from malfunction of electronic or software-controlled systems in order to derive measures how to prevent or mitigate them. This chapter explains what hazards could arise in the context of CMS and how they are systematically investigated, and what follows during the remaining development process activities from Functional Safety point of view. Although the ISO 26262 standard is a quite complex work, consisting of 10 parts, this chapter invites the reader to take a tour to the most important cornerstones of the safety process, such as Item Definition, Hazard Analysis and Risk Assessment, Functional and Technical Safety Concept, Safety Analysis, and finally the Functional Safety aspects of design and verification activities. Although an exhaustive discussion of all facets of Functional Safety is not possible on a few pages, it is important to know what activities in general need to be planned for a CMS development project, and where expert consultancy is needed. The chapter ends with an outlook of potential extensions of CMS in the future, when they may become part of more advanced assistance and warning features.

Keywords Camera monitor systems · Functional safety · ISO 26262

Abbreviations

ADAS	Advanced Driver Assistance System
ASIC	Application Specific Integrated Circuit

B. Kaiser (✉)
Berner & Mattner Systemtechnik GmbH, Munich, Germany
e-mail: Bernhard.Kaiser@berner-mattner.com

A. Terzis (ed.), *Handbook of Camera Monitor Systems*,
Augmented Vision and Reality 5, DOI 10.1007/978-3-319-29611-1_16

ASIL	Automotive Safety Integrity Level
CMS	Camera Monitor System
FMEA	Failure Modes and Effects Analysis
FMEDA	Failure Modes, Effects and Diagnostic Coverage Analysis
FPGA	Field Programmable Gate Array
FSC	Functional Safety Concept
FTA	Fault Tree Analysis
HARA	Hazard Analysis and Risk Assessment
ISO	International Organization for Standardization
SPICE	Software Process Improvement and Capability Determination
TSC	Technical Safety Concept

1 Safety Relevance of Camera Monitor Systems

At latest since the advent of the automotive safety standard ISO 26262 [1] in 2011, all carmakers and suppliers of automotive electronic systems should be aware that considering functional safety is a mandatory and integral part of a development project, just as important as assuring the system's nominal function. Even before the introduction of ISO 26262, no one would ever have put in question that vehicle braking or steering systems are safety-related systems. For systems, however, that are often described by attributes such as "informative" or "assisting", there have been and might still exist doubts about whether or not these systems require explicit safety considerations. It is often argued that these kinds of systems do not directly influence any vehicle actuators, like powertrain, steering system or brakes do. Indeed, not every electronic system in a car is equally safety critical: an air condition system is an electronic vehicle system, but obviously not as safety-critical as a steering system. In order to rank the importance for safety, the standard does consider different levels of safety criticality by providing a system of four different Automotive Safety Integrity Levels (ASILs), ranging from ASIL A (lowest) to ASIL D (highest), depending on the safety risk associated with the system. The effort required for the execution of safety tasks (e.g. formal vs. informal specification, peer reviews vs. inspection session with dedicated experts, quantitative vs. qualitative failure analysis, depth of testing etc.) increases with increasing ASILs. Actually, the ASIL is not attributed to a system as a whole, but to each individual hazard it can cause: an engine control unit, for instance, might have some ASIL D hazards, some ASIL C hazards, some ASIL B hazards and some ASIL A hazards. There is one more level below ASIL A; this one is labelled "QM" (QM stands for Quality Management, i.e. the normal quality management techniques such as ISO/TS 16949 [2] and Automotive SPICE [3] are considered sufficient for coping with these hazards). Only if all hazards of a system are ranked as "QM", we can call it a system that is not safety-related. There may be some electronic systems left in a vehicle that are actually not considered safety related. However, being an "informative" system does by no means imply that there is no safety obligation for this system (consider for example the brake lights of a vehicle: they are obviously purely informative, but their failure to indicate a deceleration is considered as a hazard, in some vehicle project ranked as high as ASIL C). It is not difficult to see that similar consideration apply to Camera Monitor Systems (CMS), as they provide necessary information to the driver, and failure to do so or the provision of falsified information can distract the driver or even cause him to perform manoeuvers that might end in an accident. So we might conclude that CMS have at least the potential to cause hazards (indirectly) and are therefore safety-relevant systems according to the ISO 26262 definition. This conclusion is fully in line with the point of view of the domain standard for CMS, the ISO 16505 [4]:

> *Camera Monitoring Systems covered by this International Standard have to be considered as safety relevant systems because*
>
> *(a) their correct function is a necessary aid to the driver in various traffic situations, and failure to perform this function can therefore lead to accidents, and*
>
> *(b) misbehaviour of those systems can lead to irritation or distraction of the driver, in consequence*
>
> *leading him to cause an accident.*
>
> *Therefore, application of the safety standards relevant to the application domain (e.g. ISO 26262) shall be considered.*

What hazards exist for a particular CMS and up to which ASIL they are ranked, will be examined during Hazard Analysis (see further down).

The safety standard ISO 26262, by the way, is in its first edition from 2011 applicable to passenger cars up to 3.5 tons only. However, it is recommendable to take it as a reference for CMS in trucks as well, because it is very likely that from the next edition it will be extended to cover trucks and motorcycles.

The term "Functional Safety" addresses only hazards caused by malfunctions of electric, electronic or programmable electronic systems. This means that the general contribution of CMS to vehicle and traffic safety is not discussed in this chapter. This also means that the ability to provide a safety operation in fault-free conditions, i.e. the nominal performance of these systems (e.g. to adapt quickly enough to varying light conditions so that the driver can perceive all relevant objects at any time) is also not in the scope of functional safety acc. to ISO 26262, and therefore left aside in this book chapter. For these aspects, the reader is referred to ISO 16505 and the current technical literature about CMS.

Once having accepted that CMS are safety-related systems, manufacturers of these systems, their sub-suppliers and finally the carmakers integrating CMS into their vehicles must all be aware of the consequences of this fact regarding the design, development, deployment, operation and maintenance of these products. The remainder of the chapter at hand guides through the main aspects added to

CMS from safety point of view, taking, of course, ISO 26262 as the most basic reference and, at the same time, restricting the meaning of functional safety to the scope delimited in this standard.

2 Particularities of CMS in Comparison to Other Safety-Related Systems

One may find carrying out the safety process for a CMS a little bit more unnatural in comparison to the safety process for a steering or braking system. First of all this might be a mere consequence of habitude: designers of steering or braking systems have been familiar with safety questions for many years before ISO 26262 even came into the world, but not so developers of cameras and displays, who often have their origins in the domain of infotainment and consumer products. But even for experienced safety specialists, CMS may bring in some interesting new questions.

The first activity in a safety lifecycle according to ISO 26262 is the Item Definition, i.e. defining the scope, the purpose, the outer interfaces and the intended functions of the item of consideration. This is in the vast majority of cases the carmaker's task, but here we encounter the first particularity of CMS: CMS are specialized and delimited systems, i.e. they have a clearly defined and commonly known task and—the most significant difference—almost no interfaces to the surrounding vehicle. The space to integrate camera, controller and display, a few mounting holes, a power supply line, that's all in most cases. This means in turn that almost every activity of the safety lifecycle can be delegated from the OEM to the CMS supplier in a specific vehicle project, or CMS suppliers can offer such systems off-the-shelf, along with all necessary safety documentation. The first case is naturally manageable in the context of ISO 26262 because this standard allows a largely flexible partition of work, which needs to be specified in a "Development Interface Agreement" (DIA) between carmaker and supplier. The second case is also explicitly supported by ISO 26262, either by making the CMS the "item" (in ISO 26262 language, see subsection about Item Definition later down) for the safety process, or, in particular cases, by applying the concept of a "Safety Element out of Context (SEooC)", which means that the system is developed under some assumptions that need to be documented in written and verified on each deployment in an actual vehicle project. Both ways are possible, and the second one is an interesting solution for CMS providers aiming at selling their solutions consisting of CMS and value-added-functions (see considerations near the end of this chapter) to different OEMs. Anyway, it should not be ignored that the operational situations and the visibility conditions in a specific vehicle are always project dependent and therefore there is no simple copy-pasting of the hazard analysis and the ASIL ratings—at least a critical review by the vehicle OEM and, if necessary, an adaptation is required.

The next particularity is that CMS in the scope of ISO 16505 (i.e. without interfacing to any driver assistance systems) cannot cause hazards directly by interfering with powertrain, braking or steering system of the vehicle, nor can they directly cause injuries to humans (as electric window-lifters do, for instance). The emergence of hazards is always via the driver, so assumptions must be made on human perception and behavior. Like conventional rear-view mirrors, CMS are an important aid for many driving manoeuvers like turning, backing up, overtaking and lane changes. But this does not mean that their failure (in particular total failure, leading to a completely black screen) in every case leads to an accident. As we will see later in the subchapter about hazard analysis, justified assumption must be made on the driver's ability to avoid a hazard (called "controllability" in terms of ISO 26262). But one misunderstanding should be avoided from the beginning: many imaginable failures of CMS can (more or less easily) be perceived[1] by the driver, but that does not automatically lead to the conclusion that these are "safe failures", not requiring any technical countermeasures, or that they are rated QM instead of an ASIL! So, when classifying failures of CMS during Hazard Analysis, the question at hand is whether or not the driver or other traffic participants can effectively take appropriate action to avoid an accident. Nevertheless it is likely that failures of CMS are the worse, the harder they are to perceive (i.e. replaying an old scenery of a moving landscape in endless loop might be worse than a totally black display). This leads to the interesting question if CMS should be considered as "fail operational" systems, i.e. systems that need to be operable all the time in order to be safe. This classification would complicate the design a lot, because power supply, processing units and even camera and display would have to be redundantly available or equipped with emergency-run functionality. Fortunately, most experts agree today that there are enough reaction possibilities for the driver, once he is aware of the failure of the CMS, at least if the vehicle is still comparable to today's passenger cars, with windows to the front and to both sides of the vehicle. In these cases, the safety tasks mostly reduce to detecting failures and presenting them in a recognizable way to the driver, trusting in his ability to control the situation.

A last difference to many other vehicle systems is the fact that many components come from other industry domains or even consumer electronics, and some of them are unfamiliar to the majority of the automotive safety engineers today (e.g. the optical parts of cameras, extensive usage of FPGAs in the signal flow, high proportion of intelligent filtering and adaptation algorithms). This may lead to unsuitable solutions or over-engineering caused by concerns due to a lack of experience with similar designs. The solution, as in any other safety-critical domain of technology is to put

[1]The ISO 26262 introduces the term of a "perceived fault", but does so in a different context (in the context of perceiving latent faults in a technical systems, that have no diagnostic mechanism, but are perceivable by the driver, e.g. by a strange noise). The term "perceived fault" must not be misused in the context of hazard classification, simply assuming that a failure of a CMS, as soon as sit is perceived by the user, is safe (a drastic example: drivers would certainly perceive it when their steering wheel is blocked due to some technical fault, but nevertheless, would have few possibilities of avoiding an accident in this situation.).

domain experts (in particular, the “normal” developers of the system, application engineers and testers) in close contact with safety experts when creating the Technical Safety Concept.

3 Interplay of Nominal Performance and Safety

So, once having accepted that CMS are safety-related systems, it is obviously *necessary* to apply the appropriate safety standard (typically ISO 26262) for the sake of safety. But is adherence to this standard also *sufficient* to make CMS acceptably safe, in all meanings of the term? Product liability legislation, for instance the German Produkthaftungsgesetz (Product Liability Act) defines a behavior of a technical system as “faulty” as soon as it does not exhibit the behavior that a user would reasonably expect. This legislation (not the scope definition of any industry standard) is taken as a reference when it comes to any lawsuits regarding product safety. ISO 26262 is a standard that copes with electric, electronic and software-controlled systems (which obviously applies to CMS), but not covering optical aspects, for instance, and, in particular, considering only hazards that are caused by *technical failures*, not by *limitations of their nominal performance* (this might change in the second edition of the ISO 26262). In other words: a hazard identified during Hazard Analysis could be that the driver cannot recognize another vehicle after a sudden change in the light conditions, because then the object does not appear on the display. If this effect is caused by a technical failure (e.g. a bug in some software function that stops the contrast adjustment algorithm working), this is subject to ISO 26262. But if it is caused by general and well-known incapability of the camera to adapt to these conditions, it is not subject to ISO 26262, although it has a similar impact on the safety of the driver. This means in consequence, that CMS manufacturers must care first for building a system to be fit for purpose, and second to care for safety issues caused by its potential failures. For the first aspect, the current legislation and state-of-the art must be observed (a central point of reference is ISO 16505) and for the second aspect, ISO 26262 must be fulfilled in addition.

A well-defined behavior of the nominal function is a precondition for proper execution of the safety analyses anyway, because failures can only be defined as deviations from nominal behavior, which must be specified to this end. When talking about “failures” in the following sub-chapters, we are referring to violations of the specified nominal behavior of the system as a whole or some of its components or functions.

Therefore, it is a mandatory part of the development of safety-relevant systems to develop the nominal function with a minimum level of process maturity. The rest of this chapter will deal with functional safety in the sense of ISO 26262 only and leave nominal performance aspects aside.

4 ISO 26262 Requirements for the Development of Safety Relevant Systems

ISO 26262 is one of the broadest industry standards ever released. It comprises 10 parts, one of them being completely informative. After a glossary of terms (Part 1), the aspects of safety process definition and safety management are described in Part 2. The next parts (Part 3–Part 7) cover the whole development and operation cycle of an automotive system. They focus on the safety-related aspects, but define them on the background of the well-known V-model process; this means in other words, that the existence of a mature development process, at least roughly following the V-model with its usual activities, is assumed as the *foundation* for the application of ISO 26262 (e.g. ISO 26262 only talks about *safety* requirements, not on requirements in general, but, of course, assumes the existence of a mature requirements management process in general). ISO 26262 clearly states that basic process maturity and the existence of a quality management system (e.g. acc. ISO/TS 16949) is the foundation for an appropriate safety process (see ISO 26262-2, Clause 5.4.4.1). A good rule of thumb is that a process maturity comparable to Automotive SPICE Level 2 is the starting point for the development of any safety-related system (corresponding to the level “QM” = Quality Management, which is the lowest level a hazard can be rated if it does not reach an ASIL A rating). Consequently, for hazards with higher ASIL rating, the ISO 26262 adds further requirements on top of this. ISO 26262 contains requirements regarding both the presence of certain activities (what has to be done) and the manner or maturity of performing them (how well does it have to be done, e.g. whether a review activity is performed as a peer review, a structured walkthrough or even an inspection with well-defined roles and a mandatory preparation). Many activities mentioned in ISO 26262 are already known from standard automotive processes (e.g. requirements analysis, testing, reviews, FMEA), some are added upon them especially from safety point of view (e.g. Hazard Analysis and Risk Assessment, Safety Concept Creation, Safety Analyses). A compact representation of the V-Model with safety additions is shown in Fig. 1.

The parts of ISO 26262 that deal with the actual development process are structured as follows: Part 3 covers the safety activities in the early project phases, in particular item definition and hazard analysis and the early phases of safety concept generation (referred to as Functional Safety Concept). Parts 4, 5, and 6 cover the development activities on system, hardware, and software level. Part 4 talks about the backbone of all safety activities, the Technical Safety Concept (unfortunately not in a distinct chapter, but rather hidden in Chap. 7 “System Design”, which also covers aspects like architecture and safety analyses). All of the three parts 4, 5, and 6 follow the same structure, starting with initiating planning activities, then going down the V-model from requirements to architecture and design steps, followed by the implementation phase and then up again the different stages of integration and verification (in particular, testing). As the development on system level is continued on hardware and software level, there is an apparent “gap”

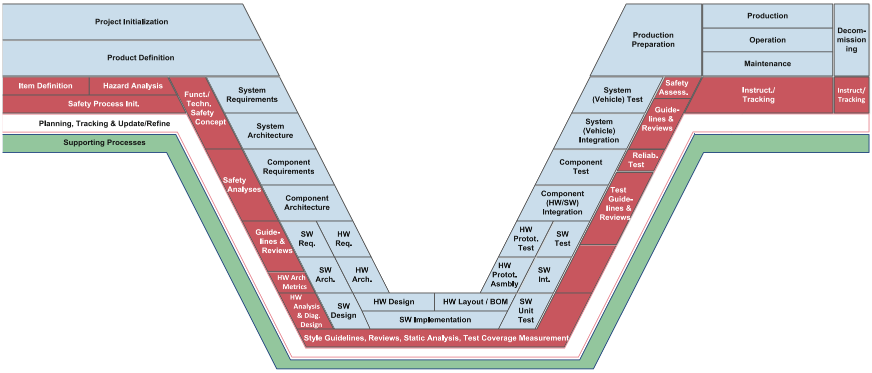


Fig. 1 V-Modell of the development process for safety-relevant systems

between Chaps. 7 and 8 in Part 4—this gap is filled by inserting the contents of Part 5 for hardware development and Part 6 for software development. The Parts 5 and 6 in detail are adapted to the specific aspects of hardware and software; in particular, Part 5 contains chapters about hardware part reliability in terms of probabilistic metrics for random failure occurrence and its impact on the hazards with respect to the given architecture. This kind of consideration (with quantitative target values that need to be fulfilled) is mandatory, with different quantitative target values, for ASIL C and D (which are often not relevant for CMS), but also recommended from ASIL B (which is a level that can be reached by some hazards of CMS, depending on the vehicle project). Whether or not these quantitative reliability considerations are required for CMS in every case will be discussed further down.

In many CMS, programmable hardware components such as FPGAs are used. They have some aspects of hardware parts and some aspects of software parts. To assure the safety of these systems, appropriate aspects from Part 5 and Part 6 must be applied.

In general, the full process displayed in ISO 26262 should be regarded as a kind of template and may (or even must) be adapted ("tailored") to the specific development project at hand, respecting:

- The nature of the product (e.g. contains HW, SW or both, maximum ASIL rating).
- The underlying development process, in particular, the agreed workshare between vehicle OEM, system supplier and sub-suppliers.

The remainder of the ISO standard (Part 8–Part 9) contain important requirements about specific activities and supporting processes (e.g. change management, requirements management) and analyses to be carried out, whereas Part 10 is purely informative and gives some additional guidance on special topics (which are recommendable to read!).

ISO 26262 is a quite complex standard and for product developers applying it for the first time, consultancy by experienced safety experts is highly recommended.

There exist also some books written by authors with automotive experience, which might be helpful for first guidance to the application of ISO 26262, e.g. [5–7].

5 Contents of an Item Definition

All safety activities refer to an “item”. An “item” in terms of ISO 26262 is defined as a “system or array of systems to implement a function at the vehicle level, to which ISO 26262 is applied”. The Item Definition marks the scope of the Safety Considerations on an overview level and is the starting point of all further safety activities. It is, in particular, a necessary preparation for the Hazard Analysis and Risk Assessment (HARA), because in order to identify malfunctions that may lead to scenarios that bear the risk of an accident (called hazards), the interfaces of the investigated system to its environment must be known, as well as the specified behavior at these interfaces. Deviations from this specified behavior constitute the item’s failures, a subset of these constituting the hazardous failures. As explained above, CMSs are well-suited to be regarded as an item according to the definition in ISO 26262. So the Item Definition usually depicts the entire CMS with camera(s), processing unit(s) and display(s).

The structure of an Item Definition is not prescribed by ISO 26262. Some companies use a commented block diagram as an Item Definition, sometimes complemented by a table listing the functions, parts or interfaces of the Item. Some use a textual document (with at least one overview drawing), or a text document (e.g. edited in a requirement management system) to structure the content. This form is the recommendable one; a text document serving as an Item Definition for a typical CMS needs not to be longer than 5 to 10 pages. It is allowed to use a document that is generated anyway during the early stages of the system development (e.g. early version of specification) as an Item Definition instead to creating an extra document.

An example of a block diagram (taken from ISO 16505) that might be part of an item definition for a CMS is shown in Fig. 2.

An Item Definition should contain at least:

- A functional block diagram with clear indication of the item’s main function blocks, its border, its interfaces and its neighbor systems and users (including driver, other traffic participants, other parts of the vehicle, interaction with the road or the traffic situation etc.).
- A brief description of the intended functionality of the item (e.g. a list of use cases or top-level requirements).
- A list of applicable standards and legislation and possible hazards already foreseeable.
- A list of assumptions under which the item is planned to be developed, integrated, and used (in case of CMS this might include questions such as: which types of vehicles, mounting positions, additional possibility for direct sideward

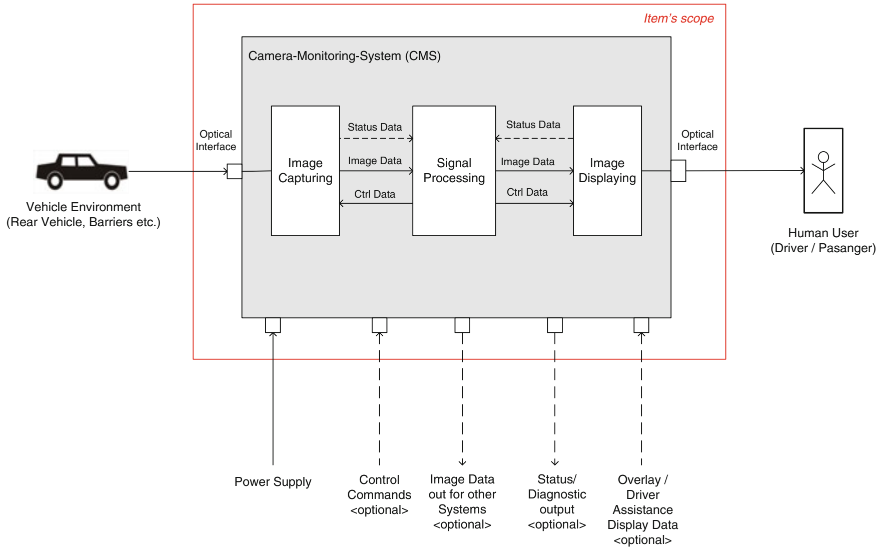


Fig. 2 Functional block diagram of a CMS

and backward sight as, right-hand vs. left-hand steering, special side conditions in case of specialized trucks, and the like).

A CMS typically has no or very few input interfaces besides the camera and the power supply. There may be operating elements like selector buttons, wheel or levers that allow the driver to select CMS functions (e.g. to re-activate it when opening the door after a longer period of sitting inside the parked car), to adapt certain parameters (within the scope that is defined by legal regulations and ISO 16505). If these are not regarded as parts of the item (e.g. if they are placed somewhere else in the car than the display, perhaps designed in responsibility of the OEM or some other supplier, and connected via busses such as CAN or LIN), they should be shown outside the item border and the corresponding data interface becomes an input interface of the Item. The same applies to other communication links, e.g. signals via CAN or other busses that indicate the vehicle speed or the direction (forward/reverse) in order to derive decisions for the CMS from it (e.g. to adapt the scaling factor or the angle of view for the camera). If such input interfaces exist, they will play a significant role later during development of the Functional Safety Concept: it will have to be shown that the input signals comply to the necessary safety integrity level (ASIL), or alternatively, it will have to be shown that even wrong input information (in any sense of this term) will not lead to a hazard.

More relevant for the HARA are the output interface of an item, as these are the interfaces where the item influences the driving behavior, issues information to the driver or other humans, or, more generally, where any kind of effect can leave the Item's border and may lead to a hazard. Usually, there is just one output interface, the display. Depending on the specific system, there might be additional output channels, for instance, status indicators or warning lamps for cases of malfunctions.

If the CMS is connected to "Advanced Driver Assistance Systems" (ADAS), the interfaces to these systems have to be shown as well, and the advices in the dedicated sub-chapter at the end of this chapter should be respected.

As the Item Definition is typically created in the initial phase of product development, it is likely that some of the contained information changes at a later point of time. Then, the Item Definition should be updated and the impact must be analyzed and necessary updates to other artefacts must be tracked by the change management process.

6 Identifying and Classifying Hazards

The next step after the Item Definition has been finished is the Hazard Analysis and Risk Assessment (HARA).

The purpose of the HARA is to find all foreseeable hazards of an item by applying a systematic search method, and to rate the hazards on a scale of Automotive Safety Integrity Levels (ASILs) from A (lowest) to D (highest), or to rate them as not specifically safety-critical (QM = normal **Q**uality **M**anagement is considered sufficient to cope with them).

The procedure can be explained a little bit by performing the steps:

(1) Identify all situations in which the vehicle is used (and that are relevant or distinguishing for the item under consideration).
(2) Identify all system malfunctions of the item in a black box view.
(3) Combine the situations with the malfunctions and investigate the potential accidents that might result.

The situation analysis (step 1) covers standard operational situations (e.g. driving in town, driving on country road, driving on highway), combined with different additional conditions relevant to the item (e.g. day, night, fog, snow, rain, tunnel), plus relevant manoeuvers (e.g. overtaking, turning right or left) and special situations, if there exist relevant ones (e.g. petrol station or carwash are usually not relevant for CMS, but opening the driver door is relevant, as the CMS is used in this case for observing the traffic from behind). Building the situation catalog requires a lot of experience, because the combination of several aspects (road type, speed, light conditions, weather conditions, objects around etc.) can lead to thousands of single cases. Justifiably building equivalence classes and subsumption relations (worst case dominates less severe cases) is mandatory to keep the task manageable. Many automotive companies have created and harmonized situation catalogs with corresponding E factor recommendations meanwhile. In any case, the standard use cases for CMS according ISO 16505, Table 2 (List of possible use cases applicable for CMS) shall be considered; only those that do not apply to the system at hand, may be left aside; a written justification should be given.[2]

The analysis of malfunctions (step 2) can be performed in FMEA style or using keywords from the HAZOP method (e.g. too low, too high, too late, unexpected…) applied to every output interface of the item. Inner faults of the item are not of interest at this stage.

The resulting hazards should be named in a recognizeable way, so that appropriate safety goals (requirements regarding the preventions of the hazards) can be derived from them (e.g. "frozen image"). If available, a safe state (e.g. "turn display off completely") should be named, and a fault tolerance time (e.g. 100 ms). The determination of a fault tolerance time may involve field studies with a significant pool of test drivers.

The ASIL for every hazard is calculated according based on:

- the occurrence frequency or relative duration of the operational situations in which a system failure could lead to an accident (not the frequency of the system failure): factor E from E1[3] (very seldom) to E4 (on almost every ride).
- the severity of the related accident regarding death or injury of humans: factor S from S0 (no injuries) to S3 (severe injuries, survival uncertain).
- the controllability, i.e. the possibility for humans (driver of own vehicle or other traffic participants) to react appropriately to prevent an accident despite the system failure: factor C from C0 (controllable without any specific reaction) to C3 (difficult or impossible to control even for experienced drivers).

The exact definitions of the individual factors can be found in ISO 26262 Part 3 (in particular, the appendices are of interest), but their application needs some experience in order to avoid misinterpretation. Attention should be paid to the E-factor, as there are two different scales from which the right one must be chosen for each scenario (and both of them may apply to CMS). An example is the manoeuver of overtaking: Regarding the situation "looking into the mirror to verify that no other car is approaching from behind", the question is "how often does this occur?", so the *frequency* interpretation shall be picked, leading to E3 rating. Regarding the situation "continuously observing a trailer during an overtaking manoeuver", the question is "how long (as a fraction of overall operation time) does this situation persist?", so the *duration* interpretation has to be chosen, leading to E2 rating. The estimation of the controllability (C-Factor) is particularly difficult in the case of CMS. This is due to the fact that the causal relationship from CMS failures to hazards is not as obvious as for, say, steering or braking systems, which directly influence the trajectory of the vehicle. The CMS just gives information to the driver and the driver decides (unconsciously in most cases) how to react. On the roads, sometimes vehicles can be seen with broken or misadjusted mirrors, but the driver

[2]Ideally, this should have been done already in the Item Definition.

[3]E0, standing for almost impossible events, is also a valid value, but seldom used in practice (e.g. aeroplane landing on a highway).

continue the journey over many kilometers, without causing an accident. Therefore, expert judgment or trial runs with different drivers in various situations, applying deliberate fault injection, are often necessary to find out the consequences and the controllability (C factor of ISO 26262) of CMS failures.

From these three factors, the ASIL is calculated according to Table 4 in ISO 26262-3 (or, easier: if the sum of S + E + C equals 7, the result is ASIL A, for 8 it is ASIL B, and so on; a sum of 6 and below leads to a QM rating).

All of these considerations are usually arranged in different columns of a table, either in a dedicated software tool for executing HARAs, or in a standard office table calculation program, such as Microsoft EXCEL. A simplified example for one failure and operational situation is given in Table 1. For layout reasons, the entries are arranged below each other; usually these entries would form the column fields of one line of a HARA table. The factors and ASIL result are illustrative only and need to be reviewed in an actual project.

For each hazard, the determining parameters should be recorded, e.g. minimum value of deviation from correct behavior, fault tolerance time (if applying) safe state (if applying), and a Safety Goal (top level safety requirement) should be assigned (in many cases it can be formulated as "Prevent < Hazard Name>").

As the HARA identifies whether or not a product is safety relevant, what are its hazards and what ASILs are allocated to the hazards (which, in turn, directs the safety process efforts to be carried out afterwards), the HARA must mandatorily be reviewed for formal correctness by an independent body (independence level I3, i.e. organizational independence).

Table 1 Derivation of a sample hazard of a CMS

Item malfunction	Unexpected zoom factor (objects appear further away than they are)
Situation	Driving on highway with several lanes at higher speed (above 100 km/h), intention of overtaking, driver looks briefly into mirror to verify that distance to vehicles arriving from behind on neighbor lane is sufficient
E-Factor	E3 (frequency)
Possible Consequence	Driver changes to neighbor lane although an approaching vehicle is close, front-rear collision at speed difference of more than 30 km/h, or pushng/forcing the other vehicle into road shoulder, leading to loss of control and subsequent collisions at high speed
	Severe or fatal injuries possible
S-Factor	S3
Possible Reaction	Driver could detect malfunction by unexpected distances between other object in display; driver will usually have a second look in the mirror or observe image for a longer time and therefore notice the other vehicle approaching faster than expected and avoid lane change
	Driver of other vehicle would observe the intention (e.g. by operated turn signal) and could avoid accident by braking or warn with horn or headlight flasher
C-Factor	C1
ASIL	ASIL A

The HARA is a process step that demands a great deal of time and effort, and it needs a lot of experience—not only to perform the HARA itself as efficiently as possible, but also to avoid over-engineering the system out of over-estimated hazards (which entails a significant increase in development cost or even unit cost), or neglecting hazards and jeopardizing the acceptance of the CMS by carmakers afterwards.

7 Typical Hazards of CMS

As the conditions of every project are different, the automotive safety community is reluctant to publishing generic lists of hazards and ASILs for any kind of system; maybe in the future there will be catalogs that are harmonized between international vehicle OEMs and system suppliers. Therefore, this book chapter does present an exhaustive list of hazards, nor an ASIL rating. Instead, the reader is given some hints what kinds of system failures and situations to think of. At the end, some example hazards with corresponding safety goal formulations are listed, but this list is not exhaustive and does not release a CMS manufacturer or OEM from performing his own HARA. For the same reasons, the hazards are represented without ASIL rating; however, in many vehicle applications, a CMS will end up in the "low ASIL category", i.e. reaching up to an ASIL of A or B.

The most obvious failures of CMS (i.e. the display being completely dark) are not the most serious ones, because they are easily perceived by the driver—but remember, however, that the fact that a failure is easily perceived does not imply that it does not cause a hazard; the relevant fact is the controllability. In many practical cases, it may be a valid assumption that most drivers, even when running on the middle lane of a highway, can react appropriately to a noticed failure of their review mirror, e.g. by very carefully changing the lanes towards the road shoulder, if necessary in combination with activating the hazard lights. This could justify a C1 rating, for example. In many other situations like turning around a corner or backing up, the driver can simply stop the car when the display recognizably fails. The more critical cases are those where the driver is "cheated" about the situation by a realistic and moving image (e.g. an old sequence running in endless loop, a significant delay, an image with unexpected zoom factor, making approaching vehicles to appear further away as they actually are). These are the cases that merit special attention.

A HARA for a CMS should consider at least the following malfunctions:

- No image (display completely dark).
- Image does not clearly display scenery according specification (e.g. adaptation to varying light conditions fails, leading to an overly dark or bright image).
- Frozen image (formerly correct image appears continuously as still image).

- Delayed image w.r.t. reality (more than specification allows).
- Wrong field of view.
- Wrong or unexpected zoom factor (e.g. wide-angle zoom factor for parking mode displayed on highway, making objects appear further away than they actually are).
- Artifacts on display (e.g. double or phantom objects, light spots, dark areas).

8 Recommended Types of Safety Analyses

Once the safety goals are known, the safety process continues by two intertwined activities:

- The Safety Analysis with the purpose of finding and rating failures that could cause some hazard.
- The Creation of the Functional and Technical Safety Concept that on the one hand defines solutions how to fulfill the safety goals, on the other hand defines mechanisms to avoid the failures identified by the safety analysis, or to detect them and react on them in an appropriate manner.

If we assume that the highest ASIL of any hazard is not higher than ASIL B, then the minimum requirement from ISO 26262 regarding safety analyses is to conduct a qualitative analysis (i.e. no need to calculate with failure probabilities) and an inductive (bottom-up) analysis. Both requirements can be fulfilled by applying a thorough FMEA [8], which has anyway been state-of-the art in automotive systems design for many years. The original table based FMEA has been extended to a failure-net based FMEA according to [9] which is recommendable for complex systems like CMS.

In addition, it is the author's recommendation to perform at least a qualitative (or just illustrative) Fault Tree Analysis [10] going down the hazards, in order to check with all relevant experts how they could arise (based on the functional system architecture) and to show where reference is made to redundant elements, which require special considerations to prove their independence. ASIL decomposition can be better justified using a Fault Tree.

A quantitative safety analysis is *highly* recommended (which, in fact, can be interpreted as "mandatory") form ASIL C, but recommended also for ASIL B. This means, it depends on the special case, and perhaps the safety regulations of the car OEM, whether or not probabilistic analysis is required. When probabilistic failures cannot be neglected (high failure rates in some parts, e.g. FPGAs or memory devices), as a compromise, at least a dedicated argument should be provided, e.g. that some Error Detecting Code (EDC) is applied and sufficient to cover the major parts of n-bit errors to be expected.

9 Recommended Content and Structure of a Safety Concept

The Safety Concept is the backbone of the safety-related development activities, as it:

(1) shows how the top level safety requirements (the Safety Goals) are systematically decomposed, satisfied with appropriate solutions and allocated to technical components of the system (hardware or software), thereby defining the safety-related parts of the system architecture.
(2) considers the failure modes identified by the safety analyses such as FMEA or Fault Tree Analysis (as long as they might lead to hazards in the sense of ISO 26262) and assigns appropriate and sufficient counter-measures to them (e.g. prevention measures at design, detection and reaction measures at runtime, process measures).

The Safety Concept unites some elements of a requirements specification (the so called Safety Requirements) and of a design specification, as it defines not only what to do, but also how to implement it technically (e.g. by requiring qualified components, by allocating safety-relevant software functions to distinct software partitions, by specifying plausibility checks, by defining redundant implementation is some cases).

The ISO 26262 mentions a Functional Safety Concept (in Part 3, i.e. Concept Phase) and a Technical Safety Concept (in Part 4, i.e. Development on System Level). To a large extent it is acceptable in practice to draw the border between both of them at will, or even unite all aspects in one single work product (which seems acceptable in particular for small, closed automotive systems like CMS). To the author's experience, however, it may bring in an additional boost in efficiency and reusability, if the functional safety concept is separated from the technical safety concept and actually restricted to *functional*, i.e. implementation-independent details, whereas the technical safety concept depends on the specific implementation in the project at hand. It is then possible to structure the technical safety concept by technical components (these may be hardware or software components). The benefit is obvious: if the implementation changes (e.g. change of microprocessor, using a FPGA instead of a microprocessor, changes in software implementation, change of sub-suppliers), the functional safety concept can be left unchanged with a high probability, and only the affected chapters of the technical safety concept need to be adapted [11].

The Functional Safety Concept (FSC) starts with a functional system architecture (function block diagram) of the nominal function, which may look similar to SysML Internal Block Diagrams, or the block diagrams used in simulation and design tools, such as the popular simulation tools Simulink from MathWorks or ASCET from ETAS. On this basis, the FSC describes how the function blocks contribute to the Safety Goals (i.e. what they must assure on their part in order to assure the Safety Goal) or can violate Safety Goals by any foreseeable misbehavior.

The step of safety requirements allocation can be performed by decomposing the safety goals (top level safety requirements) into more fine-grained requirements that can be allocated to individual function blocks. Graphical notations, such as the Goal Structuring Notation (GSN) [12], can help reviewers to understand and verify this decomposition quickly. The step of examining how potential failures of function blocks could contribute to safety goal violation can graphically be displayed by Fault Trees on high level (which may be illustrative and quite abstract at this point). Of course, these illustrative Fault Trees will have to be aligned with the results from the safety analyses, which are usually performed in parallel with the safety concept creation, but as separate activities (see previous section), feeding their results into the safety concept. An example is shown in Fig. 3, showing potential failures of some function blocks, all leading to the hazard “Frozen Image”. Without knowing much about the intended implementation, it is imaginable that failures of function blocks such as “Camera sends old image repeatedly” or “Transmission of images stops (undetected)” could lead to the top-level failure (hazardous event) “Frozen image”. Of course, the assumed functional safety modes will have to be linked to actual technical failure modes during creation of the Technical Safety Concept (TSC). The linking can be made referring to the ID that is assigned to each failure mode and each safety measure (e.g. F02 for a failure mode).

As a next step, countermeasures are defined that prevent failures from occurring or from leading to safety goal violation. If the FSC shall be kept independent from technical implementation, these countermeasures should be captured in functional, i.e. implementation-independent requirements, e.g. “It shall always be assured that…”, “It shall be prevented that…” or “If … occurs the system shall react by doing … within … ms”. Timing requirements have to be derived from the Fault Tolerance Times identified in the HARA and the required actions have to be chosen so that the system is taken to a safe state as defined in the HARA (if a safe state exists, otherwise, the requirement must be formulated with the aim of preventing the corresponding failure). Each safety requirement must be marked as safety-relevant by assigning the appropriate ASIL attribute to it. The ASIL is basically identical to the ASIL of the corresponding hazard; in cases where redundancy exists, the ASIL may be reduced according to the rules for ASIL decomposition acc. ISO 26262 Part 9, Sect. 5, but this step is in most cases

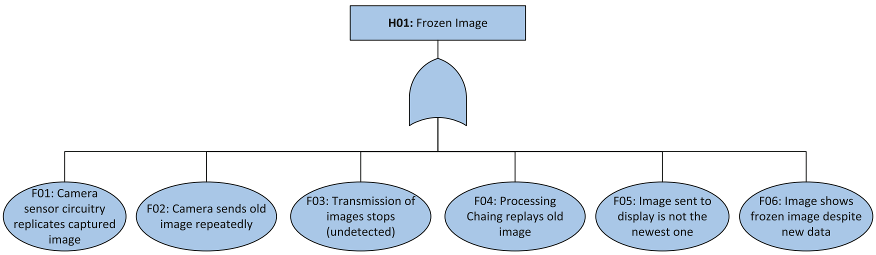


Fig. 3 Illustrative fault tree for a sample hazard of a CMS

performed on the technical level, not on the functional level. All safety requirements must be formulated acc. to the common rules for good requirements (e.g. atomicity, unambiguity etc.), which are summarized in ISO 26262 Part 8 Sect. 6. Furthermore, all functional safety requirements need to be traceable further down the process (in this case down to the technical safety concept), which can be achieved by their IDs or by technical links, if a dedicated safety tool or requirements management tool is used.

After the FSC has been created, the TSC starts. It is based on a more detailed technical system architecture, now implementation specific, which shows the actual components of the implementation (hardware and software components, data busses, digital and analog signals, and the like). The functional safety requirements are allocated to technical components, and the abstract functional failure modes are linked to technical failure modes (e.g. repeated transmission of the same image may correlate to a pointer or loop error in case of a software implementation, or to a memory addressing error in case of a FPGA implementation). If some failure mode cannot occur due to the specific realization (e.g. a frozen image on the display without actively controlling the display to do so can be excluded due to the chosen display technology), this is written down in a short argument, and technical safety mechanisms at runtime are no longer required.

The TSC part is best structured according to technical components, e.g. camera, image processing unit, controller, and display. Hierarchical decomposition into sub-components is useful if a component is too complex to be easily understood, or realized in mixed technologies or bought from different suppliers. During allocation of the safety requirements from FSC onto the technical architecture, similar requirements coming from separate hazards are harmonized and assigned the highest ASIL of all applying hazards.

In TSC, the results of FMEA should be captured and the catalog of typical hardware failures from Annex D of ISO 26262 Part 5 should be considered, in order to be sure that all potential failure modes have been dealt with. It is recommendable to use links between safety mechanisms and failure modes or to arrange them in a matrix notation (as in [11], for instance) in order to demonstrate to reviewers and assessors that all failure modes have been covered by safety mechanisms. In cases where quantitative analysis and metrics calculation is required, this demonstration has to be performed in a quantitative manner, as well, regarding the Diagnostic Coverage of each safety mechanism. This requires FMEDA or quantitative Fault Tree Analysis to be carried out. A few examples of typical technical safety mechanisms in the context of CMS could be:

- Checksums or Cyclic Redundancy Checks (CRC) in order to detect data corruption in memory or on transmission lines, and to validate the integrity of the software program or FPGA setup data.
- Timeout checks and/or rolling counters in order to detect a sudden stop or delay in a data stream, and to identify outdated image data.
- Program flow monitoring in cooperation with a watchdog, in order to detect hanging software execution (e.g. endless loops, pointer runaway).

Of course, there is much more to know about technical safety concept creation, e.g. how to care about latent faults, how to ensure freedom from interference etc.; so the recommendation for companies that are confronted with functional safety for the first time is to ask for consultancy by experienced experts. The TSC is the place where the most mistakes can occur to the author's experience.

At the end, the safety mechanisms are detailed into requirements on technical level (along with their respective ASIL attribute), their assignment to technical components (which inherit the highest ASIL of all allocated requirements and therefore need to be implemented according to an appropriate process). The intended technical solutions are explained in words and graphics. The last safety-specific activity is to export the safety requirements from the TSC to the "normal" product requirements process and to assure traceability, because otherwise there would be the risk that some requirements are overlooked and neither implemented nor tested. Once the safety requirements are captured in the requirements management system, the expectation is that designers and developers respect them for considering them in the technical product implementation, and test managers derive appropriate test cases from these requirements for subsequent verification. These are "normal" development activities; however, the ISO 26262 standard requires performing them with a certain rigor and processing maturity according to their respective ASIL. A few details on this will be presented in the remaining subchapters.

10 Design and Implementation of Software and Hardware

Regarding hardware and software design and implementation, the requirements of ISO 26262 for the lower ASILs (ASIL A and B) should most be fulfilled by a mature development process; from ASIL C onward, requirements become harder and more costly to fulfill.

When the maximum ASIL is B, for instance, hierarchical structuring of software components with restricted size of components, high cohesion and restricted coupling is highly recommended. The notation for architecture may be informal or semi-formal, which included UML or Simulink. The static and dynamic architecture needs to be described, the components have to be classified as newly developed or carry-over/third party components (which may require classification). Software error detecting mechanisms according to Table 5 in Part 6 of ISO 26262 shall be considered (e.g. control flow monitoring) and resources (CPU time, memory) shall be estimated. In the step of software detailed design and implementation, also a few common rules need to be observed, e.g. avoidance of unnecessary complexity, design for testability and maintainability and a set of coding rules (which are covered by the MISRA rules [13] that are generally recommendable for any safety-related automotive system, and which can be checked by most commercial style checkers. It is in particular the mandatory review techniques for the software detailed design and source code, that are quite time consuming even for ASIL B, because ISO 26262 requires to apply an inspection, which is a review technique with a distinct preparation phase and a review team with different roles. Deviations might be possible for the source code, if automatic checkers can prove the adherence to coding rules, the absence of certain anomaly, and the compliance with the design (e.g. by automated reverse documentation). Inspections on design level (e.g. UML activity diagrams) or model level (e.g. Simulink) is much easier performed than on C source code level.

Also for hardware, most practices required for the lower ASILs, like review of block diagrams and layout, prototype tests and environmental tests, should not be new to experienced developers. Design for testability (e.g. at end of production line) is also an issue. To work with itemized textual requirements for hardware also (as the colleagues from software development have been used to do for many years) may be a new practice for many hardware developers.

The requirements to follow in detail are much more than the few examples given above. Thus, if the development organization has no experience with the development of safety-related products, consultancy by safety experts is recommendable. Otherwise there is a risk that the product is not accepted by the OEM in the end.

11 Particularities Regarding Verification and Validation

A central part of functional safety deals with verification. Verification in the automotive domain is often interpreted as testing on different levels of software and system integration (denoted by the right leg of the V-Model), but verification is a wide range of activities covering all phases of product development. Verification activities required by ISO 26262 are, for instance, reviews for all kinds of work products (e.g. HARA, Functional/Technical Safety Concept, safety requirements, system and software architecture, software design, test cases, test outcomes etc.), which have to be conducted with a certain rigor depending on the ASIL (e.g. walkthrough vs. inspection session with distinct preparation and dedicated roles). The required activities as a function of the ASIL are specified in tables, which can be found in most chapters of the ISO 26262 standard. It is important to plan sufficient time and resources for these activities. Note that the ISO 26262 standard distinguishes between verification reviews (which should be performed by technical experts, familiar with the product and its technology) and confirmation review (to confirm that all safety process activities have been executed properly). The latter require—depending on the ASIL—persons independent from the creation process of the work products. In particular, for the HARA in every case a confirmation review with the highest degree of independence (I3, which means organizational independence) is mandatory. For small companies this might lead to the necessity to involve an external company, which is also the case for safety audits (which check the process maturity) and safety assessments (which check the overall safety of the product). In this case, the independent body could also be a person from the vehicle OEM in case of small suppliers who cannot fill this role on their own.

Besides reviews, other kinds of verification measures, such as simulations, calculations or analyses will usually apply to CMS. For aspects like microprocessor or memory load, worst case delay time for image processing or the absence of certain types of software defects, analysis or simulation techniques might be applicable to CMS. One cannot expect that all kinds of suitable verification techniques for this special class of opto-electronic systems are exhaustively mentioned in ISO 26262; nevertheless, due diligence requires to consider all suitable techniques that correspond to the state of the art for this class of systems to verify the safety-related properties. In particular, for programmable hardware devices such as FPGAs, and for certain kinds of adaptive algorithms that might apply to CMS, it is unlikely that ISO 26262 represents all applicable verification techniques.

The most prominent verification technique in automotive industries, however, is testing. Even for the lower ASIL ranges, systematic testing needs to be performed with a certain minimum depth (e.g. requirements coverage). Testing must be executed on various architectural levels (e.g. hardware testing, software unit testing, software integration testing, software black box testing, hardware/software integration testing, system testing). It is generally not sufficient to test a product only as a whole, nor is it sufficient in most cases just to show that each safety-requirement is covered by just one test case. Many requirements usually require several test cases (e.g. with different equivalence classes of input values), and for safety tests it is important to consider exception cases and fault conditions as well (e.g. by applying fault-injection testing). The variety of environmental situations may sometimes require testing techniques for these products that are not listed in ISO 26262. The ISO standard may look a bit confused sometimes with regard to the testing sections, because it lists many different approaches which are not all applicable to every kind of system, and the rating as "recommended" or "highly recommended" is sometimes given without explicit reason. To the author's opinion, a solid safety test strategy (which can be integrated with the "normal" test strategy), demonstrating coverage of all relevant testing goals, safety requirements and failure hypotheses is more convincing than a checklist that just shows that all "highly recommended" testing techniques have been applied somewhere in the process. It is important to coordinate testing (and all verification activities) between the OEM or system supplier and all sub-suppliers involved to prevent gaps in the testing strategy.

12 Additional Aspects for CMS Connected to Driver Assistance Systems

Up to now, this chapter dealt with CMS in the strict sense, being closed systems with a restricted purpose, as defined by the first edition of ISO 16505. It is not unlikely that, in the future, CMS might be extended with additional functionality and/or connected to other vehicle systems. The following cases could be imagined, for instance:

- Additional functions within the (still closed) CMS, such as changing the view angle and zoom factor for certain situations, highlighting of close objects by color, displaying distances to objects, blind spot warnings etc.
- Exporting the camera image (raw or pre-processed) of a CMS to external systems via data connections, in order to analyze the image in the external system and to provide advanced functionality (such as lane change assist, turn assist). This may comprise the generation of warnings only, or even active interventions e.g. into the steering system or braking system.
- Using the CMS display for showing additional information or warning messages, provided by external systems (e.g. highlighting an approaching vehicle, when an external radar system detects that it approaches very fast).
- Using the CMS microprocessor, FPGA or some software components shared with other functions, co-allocated to the same electronic control unit (ECU).

In any case, the considerations in this chapter regarding safety would then have to be adapted. At first, CMS are subject to legislation and standardization, so any modification of the displayed image or any display of other kinds of information must be agreed with the responsible bodies, before safety considerations even can start. It will have to be demonstrated that additional content on the screen does not impair the base function and the recognition of traffic situations by the driver, which is not a question within the scope of functional safety. Only after doing so, the functional safety analyses and the safety concept have to be extended to include the effects of potential technical failures, which could impair the function.

The absence of undesired coupling, which could lead to "cascading failures" (i.e. failure in one part causes failure in another part of the system) has to be proven in any case where new interfaces are created or functions are co-allocated onto the same technical platform (e.g. microcontroller, base software, FPGA). This property is termed "freedom from interference" in ISO 26262 language. If freedom from interference cannot be argued, all parts of the system have to be developed according the highest ASIL of any safety goal of the whole system. Interference must even be considered when the camera image of a CMS is just sent to another system via some data bus for further processing: on technical level, this might imply call to not safety-compliant bus driver functions which might in case of some software fault block program execution of the CMS core functions as well, and in consequence lead to some hazard such as a frozen image.

Another challenge might arise if raw or preprocessed images are exported to other systems for further processing: it could be the case that the functions of these systems have a higher ASIL assigned (e.g. because they might influence the steering system if they recognize an overtake attempt while detecting a vehicle approaching from behind on the neighbor lane). In this case, the exported image stream from CMS to the consuming function might be required to comply with an ASIL, say of C or D, whereas the CMS as the signal source has been initially developed to comply with ASIL A or B only. Such cases should be anticipated by

CMS designers. In any case it seems recommendable to guarantee with safety integrity (ASIL rating) that the system can detect its failures with respect to its specified performance, and not on the provision of the nominal performance itself. The consuming function would then have the responsibility to take appropriate actions and not to rely on the impaired signal. It is reasonable to prepare for these kinds of future extensions by annotating the promised quality and safety properties to any signal passed between any components within the CMS. One possible way of doing so could be the usage of contract-based development, which has gained popularity in automotive industries throughout the last few years [14–16].

13 Conclusion

Although, at first glance, Camera Monitor Systems might not be perceived as safety-related systems (in terms of functional safety acc. to ISO 26262), they actually are. Therefore, appropriate development processes to assure functional safety (i.e. the facet of safety that aims at sufficient absence of hazards and risks due to systematic and random failures in hardware and software) have to be applied by CMS manufacturers. The central standard for doing so is currently the worldwide applicable ISO 26262 standard. It assumes a general process maturity in general and specifies more specific activities, depending on the ASIL rating of the safety goals. Of course, this standard is not a replacement, but a supplement to the legislative rules and the domain specific standard ISO 16505, which have to be mandatorily considered. The ISO 26262 process is usually well applicable to CMS, because they are typical examples of software-controlled electronic automotive systems. At some points, appropriate interpretation of the ISO 26262 might be necessary. This applies in particular to the HARA, because CMS do usually not cause accidents directly, but via the information they provide (or fail to provide) to the driver. It can be expected that most CMS related safety goals (assuming that the CMS provides just the base functionality acc. to the first edition of ISO 16505) will be rated in the ASIL range between QM and ASIL B for a majority of the projects; however, it is the duty of the manufacturer to properly execute the HARA for each applicable project. The book chapter quickly summarized the main activities to be applied during CMS development because of their safety-related nature; of course, in an actual project it is mandatory to follow the ISO 26262 in deeper detail, because this book chapter cannot discuss every normative requirement. CMS offer interesting possibilities for functional extensions and combination with Advanced Driver Assistance Systems (ADAS), which have not been possible with traditional mirrors. This will of course entail new questions regarding safety, which may even increase the ASIL ratings of those systems if interventions to the powertrain, braking or steering actuators become possible in consequence.

References

1. ISO 26262 Road vehicles—Functional safety. International Organization for Standardization (2011)
2. ISO/TS 16949: Quality management systems—Particular requirements for the application of ISO 9001:2008 for automotive production and relevant service part organizations. International Organization for Standardization
3. Automotive SIG: Automotive SPICE® process reference model. URL: www.itq.ch/pdf/AutomotiveSPICE_PRM_v43.pdf. (2010). Accessed on Oct 05, 2015
4. ISO 16505: Road vehicles–ergonomic and performance aspects of camera monitor systems–requirements and test procedures. International Organization for Standardization (2015)
5. Gebhardt, V., Rieger, G.M., Mottok, J.: Funktionale Sicherheit nach ISO 26262: Ein Praxisleitfaden zur Umsetzung. dpunkt.verlag GmbH (2013)
6. Löw, P., Pabst, R., Petry, E.: Funktionale Sicherheit in der Praxis: Anwendung von DIN EN 61508 und ISO/DIS 26262 bei der Entwicklung von Serienprodukten. dpunkt Verlag (2010)
7. Ross, H.L.: Funktionale Sicherheit im Automobil: ISO 26262. Hanser Fachbuchverlag, Systemengineering auf Basis eines Sicherheitslebenszyklus und bewährten Managementsystemen (2014)
8. IEC60812: Analysis techniques for system reliability—procedure for failure mode and effect analysis (FMEA). International Electrotechnical Commission (1991)
9. VDA: Qualitätsmanagement in der Automobilindustrie Band 4: Sicherung der Qualität in der Prozesslandschaft, Kapitel: Produkt- und Prozess-FMEA. VDA (2012)
10. IEC 61025: Fault tree analysis (FTA). Beuth Verlag GmbH (2006)
11. Kaiser, B., Augustin, B., Baumann, C.: Von der Komponenten- zur Funktionsorientierten Entwicklung in der Funktionalen Sicherheit. In: 16. Internationaler Kongress Elektronik im Fahrzeug. pp. 583–594 (2013)
12. Kelly, T., Weaver, R.: The Goal Structuring Notation—A Safety Argument Notation. In: Proceedings of Dependable Systems and Networks 2004 Workshop on Assurance Cases (2004)
13. MISRA C:2012 Guidelines for the use of the C language in critical systems. MIRA
14. Bate, I., Hawkins, R., McDermid, J.: A Contract-based Approach to Designing Safe Systems. In: Proceedings of the 8th Australian Workshop on Safety Critical Systems and Software—Vol. 33. pp. 25–36. Australian Computer Society, Inc., Canberra, Australia (2003)
15. Battram, P., Kaiser, B., Weber, R.: A modular safety assurance method considering multi–aspect contracts during cyber physical system design. In: 1st International Workshop on Requirements Engineering for Self–Adaptive and Cyber–Physical Systems (RESACS). pp. 185–197 (2015)
16. Kaiser, B., Sonski, S., Buono, S., Petersen, H., Zander, J.: Lightweight Contracts for Safety-Critical Automotive Systems. In: Tagungsband 13. Workshop Automotive Software Engineering. INFORMATIK 2015—29.9.2015, Cottbus (2015)