

Doctoral Thesis

Science / Engineering

Driving Experience of an Indirect Vision Cockpit

by

Takura Yanagi

Submitted to the Graduate School of Media Design  
in partial fulfillment of the requirements for the degree of  
DOCTOR OF MEDIA DESIGN  
at the  
KEIO UNIVERSITY

Academic Year 2017



---

# Driving Experience of an Indirect Vision Cockpit

by

Takura Yanagi

Master's Degree, University of Tsukuba, 1997

B.Sc., University of Tsukuba, 1999

Submitted to the Graduate School of Media Design  
in partial fulfillment of the requirements for the degree of

DOCTOR OF MEDIA DESIGN

at the

KEIO UNIVERSITY

Academic Year 2017

Certified by .....

Hideki Sunahara  
Professor, Graduate School of Media Design  
Thesis Supervisor

Certified by .....

Hiro Kishi  
Professor, Graduate School of Media Design  
Thesis Co-Supervisor

Accepted by .....

Kouta Minamizawa  
Associate Professor, Graduate School of Media Design  
Chair, Thesis Co-Supervisor



---

# Driving Experience of an Indirect Vision Cockpit

by

Takura Yanagi

Submitted to the Graduate School of Media Design  
in partial fulfillment of the requirements for the degree of  
DOCTOR OF MEDIA DESIGN

## Abstract

Indirect vision systems that display the live video stream from a camera on a display can improve perception of the surroundings and thereby performance in a visual task compared with direct vision. For example, night vision goggles can enhance image contrast at night above the threshold needed for processing by the human visual system and enable the wearer to perceive and act on the surroundings. This thesis aims to improve the design of indirect vision systems by examining how the choice of perspective in an indirect vision system can improve spatial awareness.

Spatial awareness is the awareness of the objects around us and their location relative to each other and to our body. Different tasks may require different kinds and levels of spatial awareness. This suggests that task-specific indirect vision solutions may be useful. This thesis investigates indirect vision systems for driving cars as an example task that requires a high level of surround spatial awareness and allows implementing the necessary sensors, image processing and displays in the vehicle without requiring the user to carry the system as a mobile, wearable device.

Perspective representations, the subject of this thesis, are used to visualize three-dimensional space on two-dimensional flat displays. The components of perspective representations are the point and direction of view from which the surroundings are observed and the projection method that describes how the scenery seen from the point of view is mapped onto a two-dimensional screen. Small changes and differences in point and direction of view may be caused by binocular vision or head motion whereas large changes or differences may happen for example when moving a camera and switching from a subjective first-person perspective to a more objective third-person perspective.

This thesis aims to present a complete treatment of the subject topic by considering both, the possibility of using indirect vision to improve spatial perception beyond human capabilities of direct vision, as well as the full

---

utilization of human perception capability through indirect vision. Concerning the former, studying projection methods led to the proposal of a novel method that can present a significantly wider field of view than previous methods without degradation of distance perception important for a task like driving. How choice of point of view, known to have an effect on spatial awareness, might improve spatial awareness particularly during driving was studied in a simulated driving task. While these studies showed the potential of indirect vision to improve spatial awareness over direct vision, it is also important to make sure that indirect vision does not degrade other aspects of human perception. Motion parallax from head motion is known to be a strong depth cue but is often ignored in indirect vision systems. The final study therefore investigated whether motion parallax is needed for an indirect vision system for driving. The proposed methods were evaluated in simulated driving and using prototype implementation in real cars.

This thesis provides valuable information for the design of automotive and other indirect vision systems in the form of methods for their implementation and experiment results.

Keywords: Automotive, Driving Assistance, Virtual Reality

---

# Driving Experience of an Indirect Vision Cockpit

by

Takura Yanagi

This doctoral thesis has been examined by a Committee of Graduate School of Media Design as follows:

Thesis Reader .....  
Hiro Kishi  
Professor, Graduate School of Media Design  
Internal Examiner

Thesis Reader .....  
Kouta Minamizawa  
Associate Professor, Graduate School of Media Design  
Internal Examiner

Thesis Reader .....  
Takaaki Sugiura  
Mitsubishi Research Institute, Inc.  
External Examiner

Thesis Reader .....  
Charith Lasantha Fernando  
CTO, Telexistence, Inc.  
External Examiner

---

## Acknowledgements

This thesis would not have been possible without the unconditional support by my wife and family and the patient guidance by my supervisors Hideki Sunahara, Hiro Kishi and Kouta Minamizawa, and the support from Takaaki Sugiura and Charith Fernando. It would not have started at all without the long research relationship between Norimasa Kishi and Susumu Tachi. I am also grateful for the support from present and former colleagues at my workplace and students and staff at the University. Finally, I would like to thank the members of Nemesys Corp. and Aisan Technology Co. Ltd. with whom I had the pleasure to work together during the years that led to this thesis.



---

# Table of Contents

<b>1.</b>	<b>Introduction .....</b>	<b>1</b>
1.1	Background.....	1
1.2	Motivation.....	2
1.3	Vision.....	3
<b>2.</b>	<b>Related Work .....</b>	<b>7</b>
2.1	Human Perception and Decision Making during Driving .....	7
2.1.1	Depth Cues .....	7
2.1.2	Motion.....	9
2.1.3	Situation Awareness .....	11
2.1.4	Decision Making.....	12
2.2	Indirect Vision Systems for Cars .....	13
2.2.1	Mirrors, Rearward and Surround Camera Systems.....	13
2.2.2	Forward Camera Systems.....	17
2.2.3	Transparency .....	19
2.2.4	Cars Without Windows.....	20
2.3	Enhancing Spatial Perception in VR and Wearables .....	21
2.3.1	Expansion of Field of View .....	22
2.3.2	First vs. Third Person Perspective.....	25
2.4	Summary .....	25
<b>3.</b>	<b>Indirect Vision Cockpits .....</b>	<b>27</b>
3.1	Projection methods.....	28
3.2	Manipulating distortion to expand field of view .....	30
3.3	Manipulating point of view .....	31
3.4	Research plan .....	34
<b>4.</b>	<b>Manipulating Distortion .....</b>	<b>36</b>
4.1	Objective .....	36
4.2	Example Application.....	36
4.3	Method and implementation.....	37

---

4.3.1	Far rear center.....	38
4.3.2	Near rear sides .....	39
4.4	Experiments and Results .....	44
4.4.1	Distance perception in experiment setup.....	45
4.4.2	Experiments in real driving .....	49
<b>5.</b>	<b>Manipulating Point of View .....</b>	<b>50</b>
5.1	Objective .....	50
5.2	Which Viewpoints?.....	50
5.3	Simulator experiment.....	51
5.3.1	Experiment design.....	51
5.3.2	Experiment Task.....	53
5.3.3	Results .....	56
5.3.4	Other observations .....	69
5.4	Prototype Implementation .....	70
5.4.1	User feedback from test drives .....	75
<b>6.</b>	<b>Replicating Motion Parallax from Head Motion .....</b>	<b>76</b>
6.1	Objective .....	76
6.2	Design and Implementation .....	76
6.3	Evaluation.....	82
6.3.1	Objective and Hypothesis.....	82
6.3.2	Participants and Experiment Procedure.....	82
6.3.1	Results .....	85
<b>7.</b>	<b>Discussion .....</b>	<b>89</b>
7.1	Distortion .....	90
7.1.1	Summary.....	92
7.2	Viewpoints.....	92
7.2.1	Simulator experiment.....	92
7.2.2	Interaction of Changes in Viewpoint with Augmented Reality .....	96
7.2.1	Prototype.....	98

---

7.2.2 Summary.....	98
7.3 Motion Parallax from Head Motion .....	99
7.3.1 Summary.....	100
<b>8. Conclusion and Future Work .....</b>	<b>101</b>
8.1 Major results .....	101
8.2 Future work .....	102

---

## List of Tables

Table 4.1 : Comparison of image properties.....	42
Table 5.1 : Overview of driving tasks in the experiment .....	54
Table 5.2 : Typical driving tasks.....	55
Table 5.3 : Overview of the tasks in the experiment .....	55
Table 5.4 : Number of crashes and near misses.....	59
Table 5.5 : User feedback about viewpoint.....	75
Table 5.6 : User feedback about AR.....	75
Table B.1 : Specification of the full-HD camera.....	114
Table B.2 : Specification of the automotive camera used .....	114
Table C.3 : Experiment order by participant.....	116
Table C.4 : Experiment design.....	116
Table C.5 : Overview of objective and subjective measures for situation awareness and effort. ....	119
Table C.6 : Overview of objective and subjective measures for situation awareness and effort by experiment stage. ....	119
Table C.7 : Verbal comments.....	135

---

## List of Figures

Figure 1.1: An autonomous car moving on public roads.....	2
Figure 1.2: Conceptual image of an indirect vision cockpit in which windows are replaced by displays (right) in comparison with a conventional cockpit (left). .....	3
Figure 1.4: How indirect vision could provide a wide effective field of view on a small display within the comfort zone .....	4
Figure 1.4: How a virtual, high viewpoint reduces blindspots. ....	4
Figure 1.5: How a virtual, backward viewpoint could provide view of the rear.....	5
Figure 1.3: Examble of augmenting forward view from the driver’s seat. ....	5
Figure 1.6: Illustration of how visual experience under bad-weather condition could change.....	6
Figure 2.1: Monocular depth cues classified by their reliance on perspective. [6] .....	8
Figure 2.2: Combination of depth cues increases perception of depth. [6].....	8
Figure 2.3: Effectiveness of depth cues as a function of distance. Adapted from [3]......	9
Figure 2.4: Human perception is only good at predicting the top-most straight-line constant-speed motion.....	10
Figure 2.5: Optical flow during curve negotiation.....	10
Figure 2.6: Optical flow components (adapted from [11]).....	11
Figure 2.7: Endsley’s Model of Situation Awareness in Dynamic Decision Making. Adapted from [12]. ....	11
Figure 2.8: Rasmussen’s Skill, Rule & Knowledge (SRK) Model of Decision Making. Adapted from [13]. ....	12
Figure 2.9: Conventional convex (left) and aspheric (right) door mirrors from a press photo by Saab. [19] .....	14
Figure 2.10: Rearview camera image example.....	14
Figure 2.11: Surround view system example [20] .....	14

---

Figure 2.12: Distortion of space in surround view (left) and conventional wide-angle image (right). (Adapted from [20], axes added by the author.)	15
Figure 2.13: Framing effects	16
Figure 2.14: Visibility of the position of objects on the ground may influence distance perception	16
Figure 2.15: Electronic rear mirror [25]	17
Figure 2.16: Example of a night vision enhancement system	17
Figure 2.17: GPV Colonel 8x8x8 [27]	18
Figure 2.18: Studies of driving with indirect view, surveyed in [28]	19
Figure 2.19: Field of view and magnification levels. [28]	19
Figure 2.20: See-through A-pillars in Nissan's PIVO showcar at Tokyo Motor Show 2005	20
Figure 2.20: See-through doors and dashboard in Transparent Cockpit [29]	20
Figure 2.20: Toyota Fun-Vii showcar at Tokyo Motor Show 2011 [29]	21
Figure 2.21: Sony Concept-Cart SC-1 [30]	21
Figure 2.22: JARI ARV [31]	21
Figure 2.23: FlyViz prototype [32]	22
Figure 2.24: HMD view of FlyViz during driving on a parking lot [32]	23
Figure 2.24: The projection used in FisheyeVision (D) compared with the native fisheye camera image (A) and its undistorted counterpart (B) [33]	23
Figure 2.26: Extending field of view by blending [34]	24
Figure 2.27: Semi-transparent blending (left) and side-by-side display (right) modes of SpiderVision [34]	24
Figure 2.23: LiDARMAN [42]	25
Figure 3.1: Examples for perspective projection (left) and parallel projections (right)	29
Figure 4.2: The axially symmetric common lens model $r = F(\theta)$ is independent from the rotational angle $\beta$ .	29

---

Figure 3.2: Distortion functions of classical lens projections. [43].....	30
Figure 3.3: Illustration of an intentionally distorted view (right) providing a large effective field of view in a small display. Without such distortion, field of view is a parameter of display size. A large display might not be viewable at one glance. The challenge lies in limiting the unintended side- effects of the intended distortion.....	31
Figure 3.4: Direct vision from the driver’s head as viewpoint. .....	32
Figure 3.5: Example view from the driver’s seat.....	32
Figure 3.6: Illustration of the blindspots around a car.....	32
Figure 3.7: The Vision-by-Wire cockpit showing oneself (the orange car) within the surroundings from a virtual third person viewpoint.....	33
Figure 3.8: Location of real and virtual viewpoints.....	33
Figure 3.9: Potential of a low viewpoint to reduce close, nearby blind spots.....	33
Figure 3.10: A life-size, rearward, virtual viewpoint may be more comfortable for rearward driving than small backup camera displays.....	34
Figure 4.1: Comparison of the field of view of conventional mirrors (top; planar left door and inner rear mirror, convex passenger side door mirror on right) with that of the proposed wide-field-of-view rearview image (bottom).....	37
Figure 4.2: Predictability of motion in space when shown on a flat screen display using rectilinear projection.....	38
Figure 4.3: Illustration of the idea behind the proposed distorted projection (bottom) compared with a rectilinear projection (top).....	39
Figure 4.4: Simulated situation with a car following on the right, neighboring lane at 5m (green) and closing to 4m (red). The own vehicle is shown in gray in the center, with lines indicating the field of view of the door mirror (yellow), the inner rear mirror (orange) and the proposed electronic image (red).....	40

---

Figure 4.5: Simulated image with about 60-degree field of view with the area visible in a conventional, right door mirror marked in the center. ....	40
Figure 4.6: Simulated image with about 60-degree field of view with the area visible in a conventional inner rear mirror marked in black and the actually visible area after subtracting the blind spots from the body and rear headrests in blue. ....	41
Figure 4.7: Simulated image of the proposed projection with about 180-degree field of view from left to right edge. ....	41
Figure 4.8: Illustration of image properties in the different projections. ....	42
Figure 4.9: Extension of the axially symmetric common lens model $r = F(\theta)$ to depend on an additional parameter $\beta$ resulting in a more flexible lens model $r = F'(\theta, \beta)$ . ....	43
Figure 4.10: The proposed projection $r = F_4(\theta, \beta)$ . ....	43
Figure 4.11: Illustration of the complete system. In-camera processing means that only a camera and display are needed. ....	44
Figure 4.12: Sample images before and after processing. ....	44
Figure 4.13: A larger sample of the image after processing. ....	44
Figure 4.14: View from the driver's seat. ....	47
Figure 4.15: How the rear was visible to participants. ....	47
Figure 4.16: Experiment setup. ....	47
Figure 4.17: Detailed results. ....	48
Figure 4.18: Distance judgment error. ....	48
Figure 4.19: Interpretation of the results. ....	49
Figure 4.20: Sample images of the prototype from driving on public roads. ....	49
Figure 5.1: Possible viewpoints. ....	50
Figure 5.2: Viewpoints that can be used during forward driving as the only viewpoint. ....	51
Figure 5.4: Examples of views from the high viewpoint, normal eye point and the low viewpoint. ....	52



---

Figure 5.5: Examples of views at an intersection with AR. From left to right, high viewpoint, normal viewpoint and low viewpoint.....	53
Figure 5.7: Driving course used in the experiment.....	56
Figure 5.11: Average duration spent in each section. ....	57
Figure 5.12: Average enquete results relative to normal viewpoint on a per participant basis for all sections.....	57
Figure 5.13: Average duration of dynamic intersection without AR.....	58
Figure 5.14: Average duration of dynamic intersection with AR. ....	58
Figure 5.15: Total duration of missed opportunities. ....	59
Figure 5.16: Minimum distance from cross traffic to participant when crossing the dynamic intersection, averaged over participants and repetitions. ....	60
Figure 5.17: Range of head motion around vertical axis averaged over participants. ....	60
Figure 5.18: Enquete results for dynamic intersection without AR.....	61
Figure 5.19: Enquete results for dynamic intersection with AR. ....	61
Figure 5.20: Ratio of correct decisions. ....	62
Figure 5.21: Range of head motion around vertical axis at static intersection.....	62
Figure 5.22: Enquete scores for static intersection without AR.....	63
Figure 5.23: Enquete results for static intersection with AR. ....	63
Figure 5.24: Average distance driven off-road. ....	64
Figure 5.25: Distance driven off road during curve per participant.....	64
Figure 5.26: Distance driven off road during crank per participant. ....	65
Figure 5.27: Enquete results for curve.....	65
Figure 5.28: Enquete results for crank .....	65

---

Figure 5.29: Duration of parallel parking. ....	66
Figure 5.30: Duration of parking exit. ....	66
Figure 5.31: Number of direction changes. ....	67
Figure 5.32: Enquete results for parking exit. ....	67
Figure 5.33: Enquete results for parallel parking. ....	68
Figure 5.34: Duration of intersection by AR. ....	68
Figure 5.35: Enquete results comparing AR against no-AR. ....	69
Figure 5.41: System architecture of the prototype. ....	71
Figure 5.42: Generalized architecture adding capability to manipulate distortion in addition to point of view. ....	72
Figure 5.43: Exterior of the experiment vehicle (Virtual windows).....	72
Figure 5.44: Interior of the cockpit with the two auto- stereoscopic displays. (Virtual windows).....	72
Figure 5.45: Life-size virtual rearview with a visualization of the expected trajectory. The image is left-right reversed in order to match the visual image to the accustomed steering rotation direction. The image in the left upper corner shows the position of the experiment vehicle (silver minivan) relative to the surroundings. (Geometric enhancement) .....	72
Figure 5.46: Screen capture of the front display showing a bird view image from a high viewpoint above the car. The camera images from the surround camera have been combined with 3D geometry of the surroundings captured by the Velodyne lidar. (Geometric enhancement).....	73
Figure 5.47: The area and directions permitted to drive are colored in green (indicating area of own priority over other traffic) and orange (indicating that other traffic has priority). Three-dimensional obstacles within the road boundaries are marked in red. (Semantic enhancement).....	73
Figure 5.48: Virtual topview of an intersection with test drive course marked in green. (Combined Geometric and Semantic enhancement) .....	74
Figure 5.49: Virtual perspective birdview that combines camera image and 3D map and a CG representation of the own car. (Combined Geometric and Semantic enhancement) .....	74

---

Figure 5.50: Close perspective birdview with a transparent representation of the own car and only the wheels and steering visible to help accurate maneuvering relative to the curb on the inside of the left turn. (Combined Geometric and Semantic enhancement) .....	74
Figure 5.51: Image from a high sensitivity camera combined with augmentation of buildings, path and trajectory.....	74
Figure 6.1: System architecture. ....	78
Figure 6.2: Initial design for the Virtual Window implementation using a robotic camera in front of a large display that replaces the windshield.....	79
Figure 6.3: Side view of the system layout.....	79
Figure 6.4: Top view of the system layout.....	80
Figure 6.5: Image of the initial version of the robotic camera .....	80
Figure 6.6: Image of the cockpit interior with view to the display.....	81
Figure 6.7: Exterior of the initial version of the experiment vehicle with the robotic camera hanging downwards from the linear motor mounted on top. ....	81
Figure 6.8: Front view of the final version of the experiment vehicle with an upward robotic camera above the linear motor.....	82
Figure 6.9: View of the test course from a nearby building. ....	83
Figure 6.10: View of the test course from eye height.....	84
Figure 6.11: View of the test course from the normal car used in the base condition. The bottom of the windshield and the upper left corner were masked with black tape to limit the field of view to the same area as provided by the Vision-by-Wire implementation to ensure that the experiment results are not influenced by field-of-view and comparable.....	84
Figure 6.12: Images of an experiment run.....	85
Figure 6.13: Image of a situation where a left turn was initiated a bit too early. ....	86
Figure 6.14: The trajectories of the four experiment conditions of one of the participants. The colored lines	

---

indicate the trajectory of the center of the rear wheel. (Black dotted line: course boundary marked by white lines, red boxes: obstacles, red dotted line: stop line at the end of the course.).....	87
Figure 6.15: Enlarged top view of the trajectories at one of the left turns, approaching from the bottom right and leaving towards upper left. The colored lines indicate the center of the rear wheels. (Blue line: base condition, red: monocular, green: binocular, pink: combined binocular and motion stereo.) .....	88
Figure 8.1: Outline of future possible extensions of the method to expand field of view.....	103
Figure 8.2: Possible development of future indirect vision cars .....	104
Figure B.1: Step-by-step explanation of our image transformation.....	112
Figure B.2: System architecture.....	113
Figure B.3: Image of the camera (with a centimeter scale). .....	114
Figure C.4: HR sensor.....	117
Figure C.5: Neulog GSR sensor .....	118
Figure C.6: Time sequence of high viewpoint condition for participant 4.....	122
Figure C.7: Time sequence of normal viewpoint condition for participant 4.....	123
Figure C.8: Number of GSR peaks for participant 4. ....	124
Figure C.9: Number of GSR peaks for participant 5. ....	125
Figure C.10: Number of GSR peaks for participant 6. ....	125
Figure C.11: HR.....	126
Figure C.12: HRV data (p50NN). ....	126
Figure C.13: Total duration of missed opportunities by participant. ....	127
Figure C.14: Range of head motion around vertical axis at dynamic intersection. ....	127
Figure C.15: Number of direction changes during parking exit.....	128

---

Figure C.16: Number of direction changes during parallel parking. ....	128
Figure C.17: Duration of static intersection by AR. ....	129
Figure C.18: Duration of dynamic intersection by AR. ....	129
Figure C.19: Relative enquete score comparing AR against no-AR by participant. ....	130
Figure C.20: Relative enquete score comparing AR against no-AR by participant. ....	130
Figure C.21: Relative enquete score comparing AR against no-AR by participant. ....	130
Figure C.22: Relative enquete score comparing AR against no-AR by participant. ....	131
Figure C.23: Relative enquete score comparing AR against no-AR by participant. ....	131
Figure C.24: Relative enquete score comparing AR against no-AR by participant. ....	131

---

---

# 1. Introduction

Indirect vision systems that display the live video stream from a camera on a display when completing a visual task can improve human performance over direct vision. For example, night vision goggles can raise image contrast in darkness at night above the threshold needed for processing the visual information by the human visual system. This thesis aims to improve the design of indirect vision systems by examining how the perspective used can improve spatial awareness.

Spatial awareness is the awareness of the objects around us and their location relative to each other and to our body. Different tasks may require different kinds and levels of spatial awareness. Sitting at a desk working on a computer will require only minimal spatial awareness, like the location of the keyboard, the mouse and the display. On the other hand, playing soccer on a soccer field will require a much higher level of spatial awareness, like the own, team mates', opponents' and balls location, direction and speed on the field as well as the location of the goals.

This thesis investigates indirect vision systems using car driving as an example application that requires a high level of surround spatial awareness like the geometry of the road ahead, the longitudinal and lateral location of the own vehicle within the current lane, and the location and movement of surrounding cars, bikes and pedestrians. This does neither mean that the methods developed, nor the results obtained are limited to cars and car driving. For example, once sensors become smaller and headset-style displays more powerful, a wearable version of the proposed system could be worn 24/7.

## 1.1 Background

Cars are an important tool for mankind to support modern life by fulfilling a large percentage of transportation and mobility needs. According to official statistics, in Japan, a country with a population of about 120 million, about 80 million registered vehicles travel 500 billion kilometers per year. Cars already combine mobility enhancing technologies like the drivetrain and chassis with perception enhancing technologies like headlights, wipers, camera systems and navigation. Integrating an indirect vision system into

---

a car is therefore in some way a smaller step than making a wearable version consisting of a computer, cameras and displays.



Figure 1.1: An autonomous car moving on public roads. [1]

## 1.2 Motivation

In most countries, regulation for cars are guided by the 1968 Vienna Convention on Road Traffic [2], which stipulates that a driver is always in control of the vehicle and responsible for its behavior in traffic. The driver has to look around, understand the surrounding situation, make decisions and properly operate the vehicle to execute the intended movements. Considering the complexity of this process, driving is surprisingly safe, yet accidents occur.

The three most frequent accident causes in Japan according to official statistics are (1) distraction, (2) speeding and (3) human failure in understanding the surrounding situation, decision making or operation of the vehicle. Such human failure may have several causes. Human eyes have a limited field of view, resolution, sensitivity, location and direction. The human brain works with only a single focus of attention, which means that attention will be allocated to multiple targets only sequentially one after the other. The brain is also limited in the size of its memory, susceptible to false reasoning and prone to bad customs.

Will it be possible to overcome these limitations of human perception by manipulating the visual information that the eyes receive, and if so, how? And will it be possible to apply the findings and methods to tasks other than driving?



---

### 1.3 Vision

An indirect vision system that replaces the windows of a car by a combination of cameras, other sensors and displays would give the developer large power – and responsibility – to modify the visual perception of the driver. Output from multiple sensors at multiple locations, maybe even some of which are not mounted to the car but on other surrounding cars on the road infrastructure, could be combined to provide a complete, blindspot-free model of the surroundings, which could then be streamed to the display from any viewpoint. That image could further be augmented with distance, speed, driving rules and other useful information, and then finally be shown to the driver in place of the view through the windows in a conventional car.

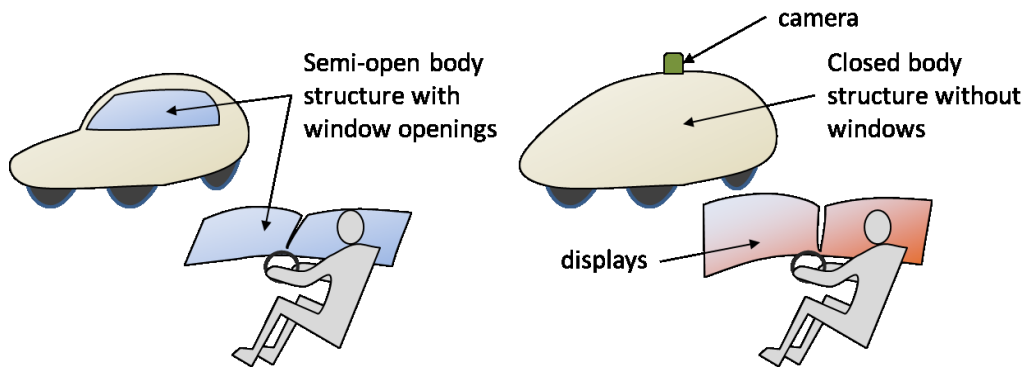


Figure 1.2: Conceptual image of an indirect vision cockpit in which windows are replaced by displays (right) in comparison with a conventional cockpit (left).

There are many possibilities how a human driver could benefit from such a substitution of direct vision through windows:

- While the total horizontal field of view of human perception is about 180 degrees, only a narrow subset in front of up to about 60 degrees is usable for symbol recognition. If the visual information from a wide field of view could be compressed and presented in a smaller field of view, humans may find it easier to achieve a better spatial awareness of the surroundings than with direct vision.

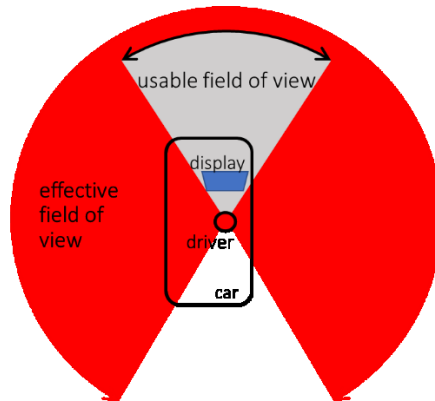


Figure 1.3: How indirect vision could provide a wide effective field of view on a small display within the comfort zone

- Other cars and pedestrians surrounding the car could be shown from an overhead viewpoint, in three dimensions and in real life size compared with the small view in current systems.

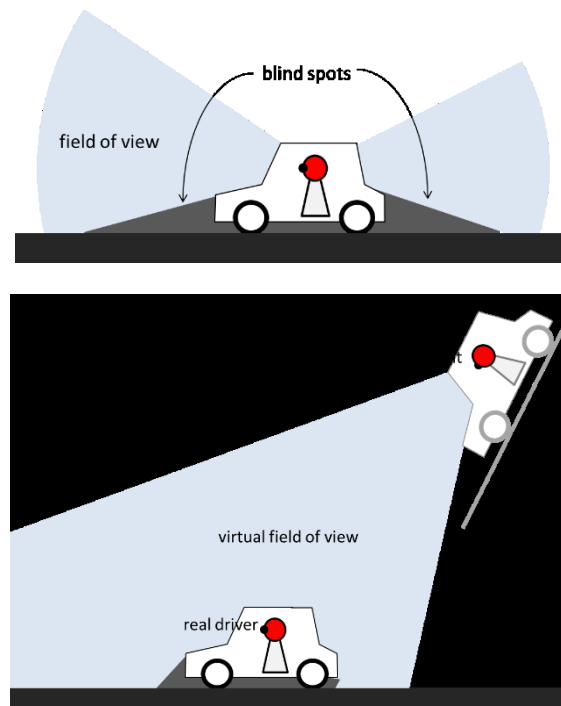


Figure 1.4: How a virtual, high viewpoint reduces blindspots.

- The rearward view could be displayed in life size in front of the driver when backing up. The driver would no longer need to turn the head to look backward, to imagine the rear situation from narrow views provided by mirrors, or to observe the image of a rearward camera on the small display of a car navigation system.

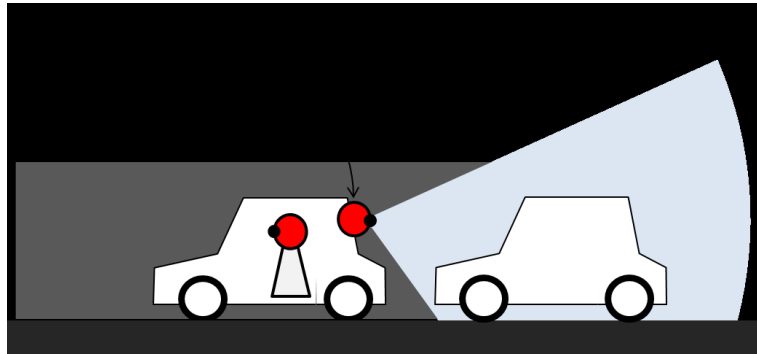


Figure 1.5: How a virtual, backward viewpoint could provide view of the rear.

- The shape of the road could be displayed beyond corners and through buildings and other occluders like trucks, for example potentially reducing the risk of misjudging safe driving speed.



Figure 1.6: Example of augmenting forward view from the driver's seat.

- A bright, real-size night vision image could make night time driving easier than seeing through windows with the aid of head lights or using conventional night vision systems which have only a very narrow field of view. Similarly, a virtual good-weather image of the surroundings could be displayed when driving in rain, making it much easier to see for example the lane markers.



Figure 1.7: Illustration of how visual experience under bad-weather condition could change.

Electronic vision could be beneficial not just for manual driving but also in a shared collaborative driving setting between the human driver and an intelligent car, ensuring that the driver sees the same information that the car sees, making it less likely that driver and car will fight with each other. It could also benefit the visual experience of riding as a passenger in an autonomous car by the ability to watch either an enhanced or virtual scenery while enjoying the privacy of a windowless vehicle without the feeling of being locked inside a claustrophobically small box.

In most countries, under current legislation, forward and sideways driver visibility must be direct and cannot be realized by displays. But as autonomous cars are in some places no longer required to allow for driver intervention and can be designed without a steering wheel, it is not unthinkable that a driver-operated car could be realized with electronic, indirect vision if it was equipped with an autonomous system as a safeguard and backup.

Once the methods and technologies are established, they could be applied to other domains. A wearable system built using similar or adapted methods could aid when walking or when using other forms of transportation like bicycles. Other systems could target spatial awareness when doing stationary tasks that require reacting to the surroundings.

---

## 2. Related Work

### 2.1 Human Perception and Decision Making during Driving

Vision is the major source of information during driving. It has been said that more than 90% of information for driving is visual. [2]

Human vision is one of the better studied functions of the human brain. It has evolved to recognize objects and classify them as, for example, cars, pedestrians or bicycles, estimate their location and motion within the static environment like roads, or to project when a moving and a static object or two moving objects will crash. Human vision is also capable of recognizing and interpreting symbols like lane markers, traffic signs and traffic lights. Most countries have rules for the visibility of the surroundings from the eye point of the driver specifying the area of windows or the number, size, location and curvature of mirrors. Rules might also specify visual acuity and color vision requirements for the driver of a vehicle.

#### 2.1.1 Depth Cues

Depth perception, i.e. the perception of the distance to an object, is essential for making decisions during driving and operating the controls of the car. The human visual system utilizes several different cues.

- Monocular depth cues are 3-dimensional interpretations obtained from a single 2D image. Monocular cues can be further divided into perspective cues, like occlusion, relative and familiar size, shadows, location of objects on the ground between the point-of-view and the horizon and other cues like shading, atmosphere and focus. [3] Stewart et al. point out that time-to-collision is often misperceived if the pedestrian is a child because of size. [4]
- Oculomotor depth cues are cues from the inward movement of the eye-balls towards a single location in space (convergence) and focusing of the eyes at a distance in space (accommodation). Oculomotor cues are considered important for immersive viewing, but their effect is difficult to measure.
- Binocular depth cues or binocular stereo combines the information from the left and right eye in the center of the visual field that is visible from both eyes (approximately 120 degrees out of about 200 degrees) and said to contribute to a sense of presence.
- Depth from motion or motion stereo uses the change of objects in the visual field, either from motion of the viewer of the viewed object to make

assumptions about the depth. A recent driving simulator study failed to identify effect of motion stereo. [5]

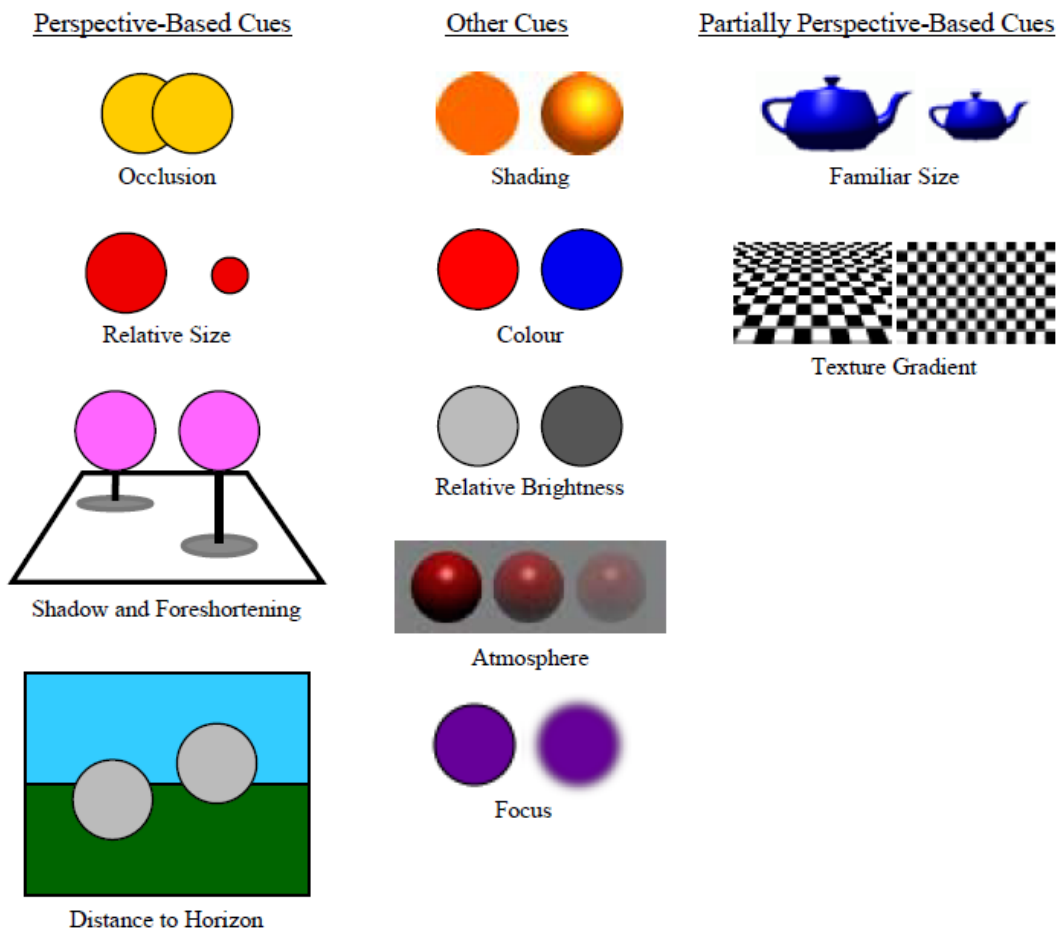


Figure 2.1: Monocular depth cues classified by their reliance on perspective. [6]

The combination of multiple depth cues can intensify the perception of depth. The cues interact with each other, making it difficult to isolate the contribution of specific cues.

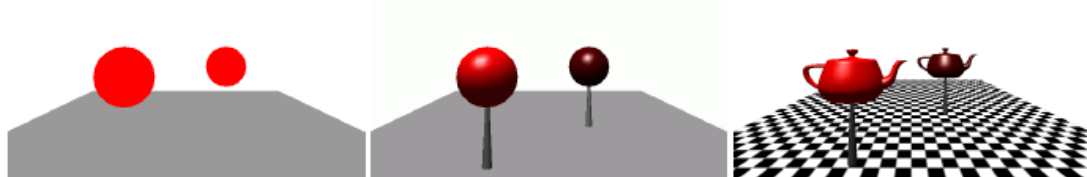


Figure 2.2: Combination of depth cues increases perception of depth. [6]

Nevertheless, [3] attempted to compare the effectiveness of cues depending on viewing distance and relative to each other. While this may provide a rough guidance when designing an indirect vision system, in reality, the

---

strength of each cue may differ depending on the actual visual properties of the situation.

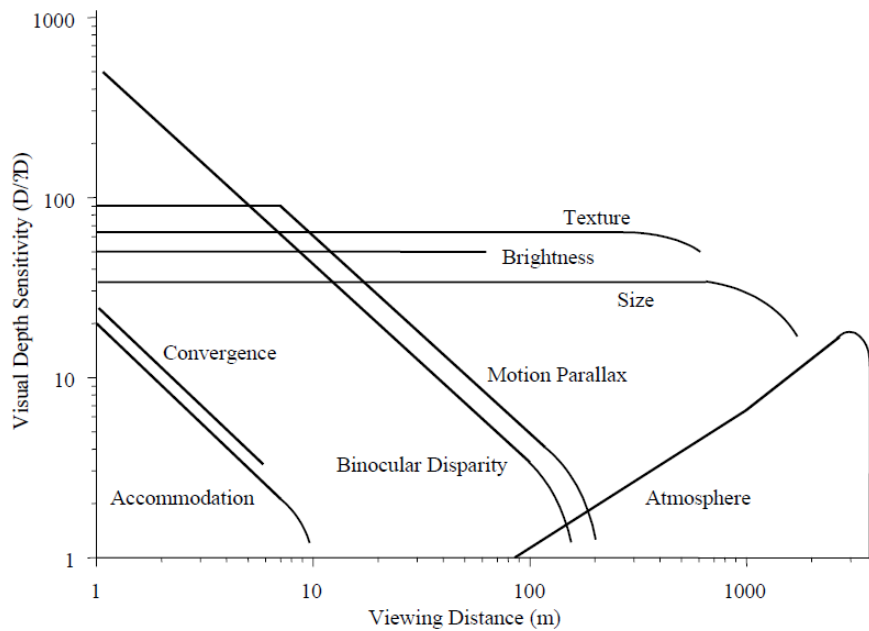


Figure 2.3: Effectiveness of depth cues as a function of distance. Adapted from [3].

For slowly avoiding obstacles in a narrow road, the most used range might be between about 2 and 15 meters making binocular disparity and motion parallax the most effective cues, whereas the range used while driving fast on a multi-lane highway could be between about 20 and 100 meters, making texture and brightness the most effective cues. In some situations, less effective cues may override more effective ones and confuse drivers.

### 2.1.2 Motion

Depth by itself may not be that useful during driving. To anticipate and avoid possible collisions, drivers must judge direction and time-to-contact. While human perception is good at predicting motion and time-to-contact if speed (first order) and direction are constant, acceleration (second order) and changes in direction make prediction difficult. [7] Binocular information can be useful but may not be reliably usable for example in mirrors. Humans may therefore use shortcuts like framing effects for practical judgments, like deciding to decelerate as soon as the engine hood of the own car hides the ground between the own car and the car in front. [8]

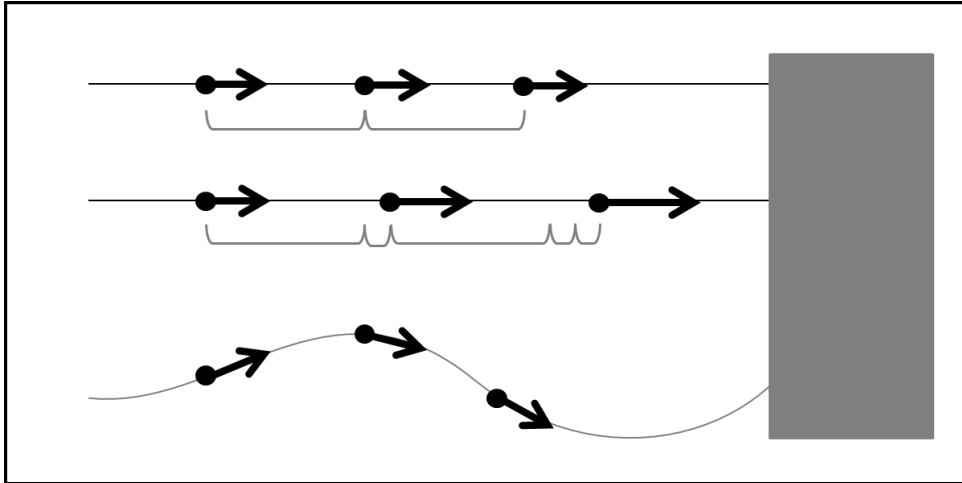


Figure 2.4: Human perception is only good at predicting the top-most straight-line constant-speed motion.

Human vision has been shown to utilize cues from optical flow to perceive the own translational and rotational movement in space relative to the environment. This is used for controlling speed and negotiating curves. Optical flow can be perceived by fine elements like detail textures in the fovea or by large elements like terrain structures in the peripheral vision. Having a large field of view is therefore considered necessary for accurate perception of ego-motion. [9]



Figure 2.5: Optical flow during curve negotiation

Optical flow changes depending on the surrounding environment, the driving style and the outward visibility from the driver's seat and can in return influence the way of driving. A higher point of view, for example, can lead to faster driving, probably caused by the reduction of optical flow in the visual field. [10]

The direction of optical flow caused by a static background during driving can be decomposed into vertically rotational, horizontally rotational, circular and radial components, respectively caused by pitching, direction



changes, rolling and longitudinal motion. The speed will depend on the distance to object in view. When driving in a straight line, the only direction present is radial flow.

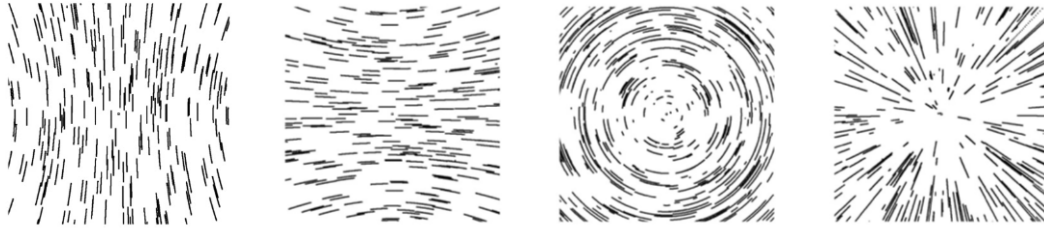


Figure 2.6: Optical flow components. Adapted from [11].

### 2.1.3 Situation Awareness

The result of perception is sometimes described by a construct named Situation Awareness (SA). It describes the “perception of the elements in the environment within a volume of time and space,” referred to as level 1, “the comprehension of their meaning” (level 2), and “the projection of their status in the near future” (level 3). [12] Spatial awareness, defined as the awareness of the objects around us and their location relative to each other and to our body, is roughly similar to SA level 1 but typically has a stronger emphasis on awareness of oneself within the surroundings.

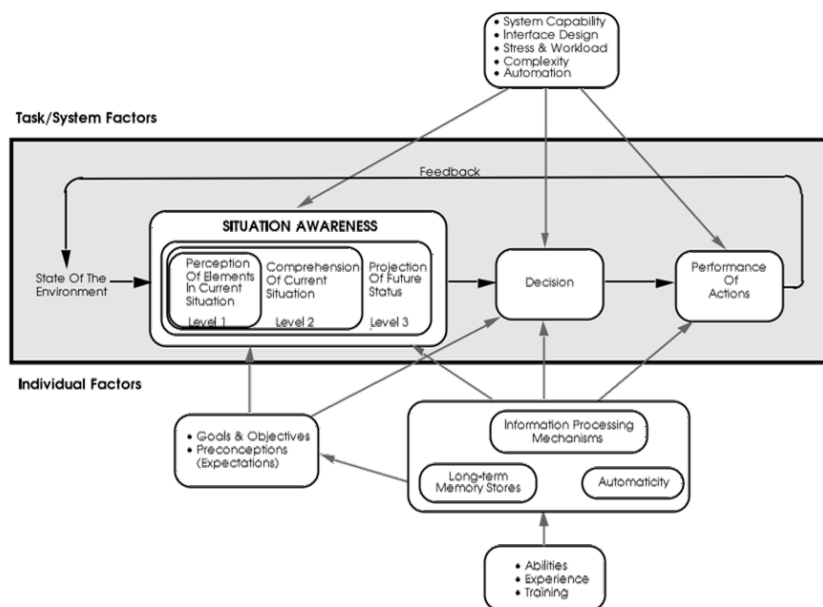


Figure 2.7: Endsley’s Model of Situation Awareness in Dynamic Decision Making. Adapted from [12].

## 2.1.4 Decision Making

Cognitive science classifies the driving task as a cognitive decision-making process. An often-cited model of decision-making in the context of driving is Rasmussen's Skill, Rule & Knowledge (SRK) Model. According to this model, the decision-making process is a cycle taking input from sensory perception of the surrounding situation like road geometry, traffic rules, obstacles and other cars through sensory information, mainly through vision, but also auditory and haptic channels. The decision-making process provides output in form of the driver operating the car by turning the steering wheel or pressing either the brake or the acceleration pedal to correct the trajectory of the car. Between input and output, there can be

1. a simple, often unconscious application of a practiced skill, like pressing the brake pedal with the right foot at exactly the correct pressure to halt at the stop line ahead, or making tender course corrections to smoothly follow a curve,
2. a pattern-based rule application that then leads to a practiced skill application, like noticing a red traffic light and initiating deceleration, or noticing a slower moving truck ahead and making a lane change to pass it,
3. or a knowledge-based decision-making step to deal with a less common situation, that is then followed by lower level actions, like deciding which of several possible routes to take depending on current surrounding traffic flow and experience.

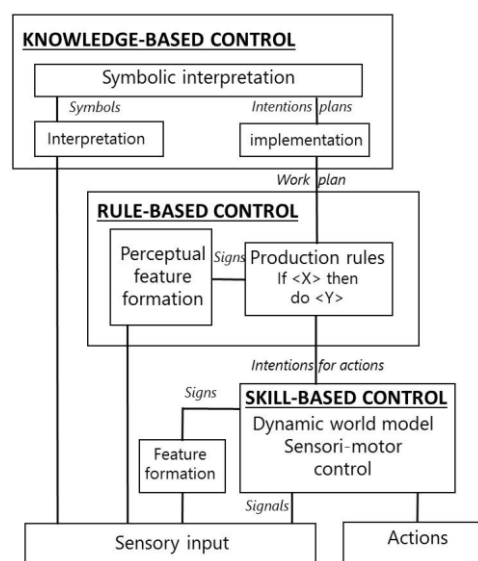


Figure 2.8: Rasmussen's Skill, Rule & Knowledge (SRK) Model of Decision Making. Adapted from [13].

---

## 2.2 Indirect Vision Systems for Cars

### 2.2.1 Mirrors, Rearward and Surround Camera Systems

Mirrors that utilize a reflective surface to change viewing direction and viewpoint are the most basic method providing indirect vision. The Vienna Convention stipulates that cars have rearview mirrors that enable the driver to see the traffic in the rear. Actual requirements differ between countries but the combination of inner rear mirror and door mirrors on each side of the vehicle have become prevalent. Larger vehicles like SUVs and trucks are often required to have additional curved mirrors to reduce nearby blind-spots. The recent increase in mirrors might partially have been an effect of the decrease in direct visibility due to aerodynamic, design and safety developments. The mean horizontal fields of view of left (driver-side), center, and right (passenger-side) mirrors of a selection of US passenger cars were reported in [14] to be 12.9, 25.3 and 22.5 degrees respectively.

One important property of mirrors is their effect on distance perception. In [15], Hecht and Brauer compared planar mirrors, that show objects in real size, with non-planar mirrors, and showed that the former provides more reliable distance perception than the latter. Interestingly, perception of objects through mirrors is complicated [16], and clarifying how different factors like viewing size, perspective cues, binocular disparity, framing effect and many others contribute to such differences is not easy. Although planar mirrors are intuitive, other mirrors may be preferable in real driving. For example, De Vos has suggested in [17] that non-planar door mirrors provide higher degrees of situation awareness than planar mirrors.

Mirrors with a large field of view often have some distortion. In [18], Hicks and Perline present a unique distortion-free mirror for rear visibility that covers a relatively wide field of view of about 45 degrees compared with typically less than 20 degrees in conventional driver-side door mirrors.

Aspheric door mirrors, show the rear around the vanishing point at a size close to uniformly convex mirrors, but add a distorted, horizontally compressed image at the outer edge. While the usefulness of aspheric mirrors to increase situational awareness has been established in human factors studies, such mirrors still have a view of less than 50 degrees and the high distortion in the outer areas only allow for checking the presence of obstacles.



Figure 2.9: Conventional convex (left) and aspheric (right) door mirrors from a press photo by Saab. [19]

Recently, electronic systems combining cameras and displays to assist driver vision have quickly become widespread in cars and are expected to further increase. Rear view cameras eliminate a large part of rearward blindspots and will be obligatory for new cars in the US soon.

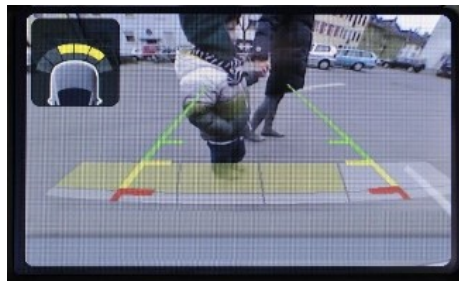
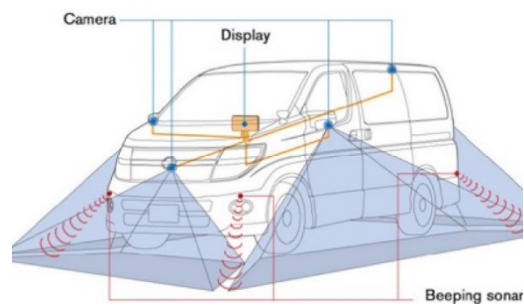


Figure 2.10: Rearview camera image example

Cars equipped with surround view systems that present the surroundings from a virtual overhead viewpoint are now available from several manufacturers.



Topview image example (left)



System overview

Figure 2.11: Surround view system example [20]

While maneuvering a nonholonomic vehicle operating its controls can still be demanding, surround view systems may have multiple benefits for perception.

- They show the surroundings that otherwise require multiple glances in different mirrors or multiple head turns in one continuous image requiring only one glance with only scanning a small display.

- They show a distortion free image of the ground plane making relative motion towards obstacles easier to predict than either distorted images and in direct view of the surroundings.
- The body of the own vehicle does not cause blind spots.



Figure 2.12: Distortion of space in surround view (left) and conventional wide-angle image (right). (Adapted from [20], axes added by the author.)

Other systems are a more direct replacement of conventional ones. Rear visibility in cars has traditionally been realized by mirrors. Mirrors provide new viewing directions for example to the rear or to blindspots from the eye point of the driver and without having to turn around. Vision by mirrors is limited by geometry for example in the direction of view, the size and the viewpoint they can provide. Electronic inside rear mirrors are available for several car models and replacing reflective door mirrors by electronic ones have been legalized in Europe. Reflective and electronic mirrors both require to take the eyes off the road in the front.

In [21], Flannagan and Sivak argue that a camera-based system with a single display location may reduce driver workload when compared with a system in which the driver must distribute attention to multiple locations. In [22], Flannagan, Sivak and Simpson argue that the lack of binocular distance information in 2D displays when compared to mirrors is not a fundamental problem. Factors like image magnification, camera location and direction, existence/absence of 3-dimensional cues or aspect ratio and size of the display as well as the expectations and habits of the participants

---

interconnect, making it difficult to reach a universal conclusion for even simple issues like the best magnification. While it may seem optimal to simulate a flat mirror, [23] showed that distance is underestimated in 2D displays possibly due to distance perception being influenced by the size of the objects relative to the size of the display which they called framing effects. This effect was not observed in mirrors. Note though that frames can be helpful for quickly judging distance relative to a threshold represented by the frame and viewers often complain about lack of orientation when looking at a wide-angle display without a frame near the region of interest. The bottom frame can influence distance perception by hiding the position of the object on the ground which is a useful perspective cue. Anecdotally, having a part of the own car visible in the mirrors helps with orientation and distance judgment but sound reasoning and scientific validation seems lacking. Most camera systems also trade off correct distance perception against field of view. Either drivers can judge distance reliably or see a wide area in one glance, but not both. [24]



Figure 2.13: Framing effects



Figure 2.14: Visibility of the position of objects on the ground may influence distance perception

While such systems partly alleviate the limitation of human situation awareness caused by the location, direction or sensitivity of the eyes, they require shifting the eyes to a display and focusing on the image shown in the display. This reduces the time that the eyes of the driver are oriented towards the direction of movement, thereby making it difficult to avoid sudden obstacles that require immediate reaction.



Figure 2.15: Electronic rear mirror [25]

### 2.2.2 Forward Camera Systems

Cars have been equipped with lights to improve visibility of the surroundings as well as the visibility of oneself to others since early on in automotive history. Headlights provide enough light to drive on dark country roads without any road lighting. Road lighting is preferable in places with dense populations as lighting the path of a vehicle with headlights mounted at a low height is fundamentally limited by geometry, for example by exponential decrease in brightness with distance, low contrast in the lit area due to front lighting, shadows from uneven road surface and the potential to blind others. Some of these issues are addressed with Night Vision Enhancement Systems (NVES) that cover a narrow but longer distance range without blinding others.

Some evaluations of NVES observed that detection performance improves at a higher overall workload which can be mitigated by adding auditory warnings but not visual augmentation. [26]



Figure 2.16: Example of a night vision enhancement system

---

Manned military vehicles are sometimes equipped with indirect vision systems that can be used in place of direct vision when the window openings are hidden under armour. Opinions differ about the presence and amount of discomfort and contributing factors like motion sickness. Improper geometry of the visual representation as well as contradicting information from the visual versus vestibular and kinesthetic channels, caused by delay of the displayed image and offset of camera viewpoint from the eyes may lead to motion sickness. While most systems are monocular, a few researchers have investigated stereoscopic systems (with fixed cameras and without motion parallax) and concluded that stereoscopic systems have an advantage in depth perception resulting in higher task performance, while having no significant effect in reducing motion sickness. [6]



Figure 2.17: GPV Colonel 8x8x8 [27]

In a review of human factors literature between 1986 and 2001 relevant for design of Night Vision Enhancement Systems, Tsimhoni and Green surveyed research about indirect vision systems for driving. [28] Some of the research used a given display size in combination with different lenses, possibly aimed at identifying the best combination of available alternatives, meaning that field of view and magnification were not independently controlled. Life-size presentation (magnification = 1.0) was often recommended, likely because of intuitiveness for depth judgments and control, but depending on task or field of view, a wider field of view with smaller magnification was preferred. At that time, and possibly limited by freedom of layout in military vehicles, no investigation replicated life-size view with a wide field of view similar to direct vision using large displays. There also was no study investigating the effect of depth cues from head motion. Studies often measured driving performance, workload and preference, but none attempted to measure user experience.



Independent measures	Dependent measures			
	Distance / gap estimation	Driving performance	Subjective workload and preference	Motion sickness
Field of view and magnification	Brown '86 Conchillo '96	Glumm '92 Oving '01 Padmos '96 Smyth '01 Sudarsan '97 van Erp '97 '98 '99	Glumm '92 Smyth '01 van Erp '99	Glumm '92 Oving '01
Camera position		Glumm '97 Padmos '96 Smith '70 van Erp '98	Glumm '97	
Stereoscopic or monoscopic	Holzhausen '93	Drascic '91 Holzhausen '93 van Erp '99	van Erp '99	
Color or monochromatic	Miller '88a '88b		Miller '88a '88b	
Panning camera	Miller '88a	van Erp '97	Miller '88a '88b	
With or without a camera		Oving '01 Padmos '96		
Image quality: Resolution, frame rate, image delay,	van Erp '98	Sudarsan '97 van Erp '98	van Erp '98	
General discussion	Padmos '95			

Figure 2.18: Studies of driving with indirect view, surveyed in [28]

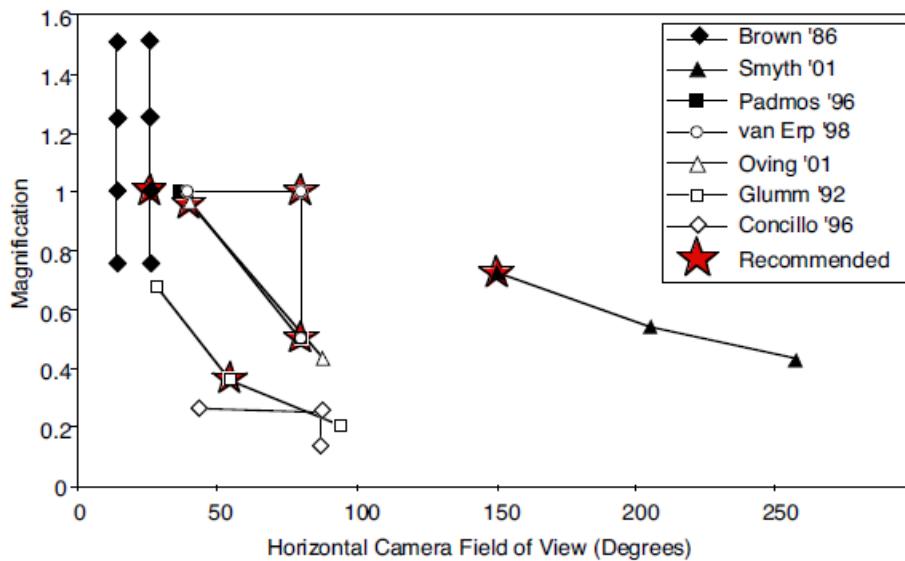


Figure 2.19: Field of view and magnification levels. [28]

### 2.2.3 Transparency

Occlusion caused by the body of the car, the occupants or luggage as well as on-road and off-road objects limit visibility during driving. Simple systems may show the video stream from a camera on the outside to a display mounted on the inside as has been done for example in Nissan's PIVO show-car (Figure 2.20). Transparent Cockpit [29] utilizes retro-reflective

---

projection to utilize not only plane surfaces but also more complex shaped interior parts as a screen which allows larger display size. While both systems succeed in eliminating or at least reducing blind spots, they do not reconstruct the surroundings from the driver's viewpoint, causing slight misalignments especially for nearby objects. These misalignments, together with the invisibility of the border of the own car may make it difficult to accurately judge distances to obstacles when used for maneuvering and not just detecting obstacles in the blind spot.



Figure 2.20: See-through A-pillars in Nissan's PIVO showcar at Tokyo Motor Show 2005



Figure 2.21: See-through doors and dashboard in Transparent Cockpit [29]

#### 2.2.4 Cars Without Windows

Some proposals for cars without windows have been made in the past. In the domain of concept cars, some examples are Toyota Fun-Vii (2011), Mercedes F105 Luxury in Motion Concept Car (2015) and Sony SC-1 (2017). System details and evaluation results have not been made publicly available.

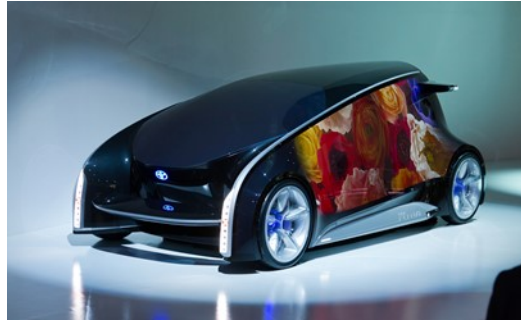


Figure 2.22: Toyota Fun-Vii showcar at Tokyo Motor Show 2011 [29]



Figure 2.23: Sony Concept-Cart SC-1 [30]

In research and education, driving simulators can be considered a kind of cars without windows. JARI developed a car with three, large displays placed in front of the driver for studying driver behavior on a closed test course using mixed reality.



Figure 2.24: JARI ARV [31]

### 2.3 Enhancing Spatial Perception in VR and Wearables

Virtual Reality (VR) aims to create a realistic visual, auditory or other experience of a virtual environment. A subfield of VR has investigated methods that tweak the realistic representation of the environment, for example, to improve perception under the constraints of the limited field of view in

---

available HMDs or task performance beyond the limitations of human perception. Two main directions will be explained in detail.

### 2.3.1 Expansion of Field of View

Special eyeglasses equipped with prisms have been proposed to expand the field of view of people with reduced peripheral vision. These usually require intensive field training. Video see-through (VST) displays used for VR are more complex but also more powerful and several proposals have been made for expanding the field of view in an intuitive way. Cramming a large effective field of view into a limited geometric field of view may come at a cost. Basic approaches are minification, potentially at the cost of size and distance perception, distortion, potentially at the cost of natural spatial perception and amplification of head motion, potentially at the cost of motion sickness.

Flyviz [33] uses a HMD to display the 360-degree image stream captured by a head mounted catadioptric camera. While it can display an integrated view of the external environment and the interior during driving as shown in Figure 2.26: HMD view of FlyViz during driving on a parking lot Figure 2.26 and driving explicitly stated as a usage scenario, the equirectangular projection used to map the 360-degree field of view of the camera onto the limited geometric field of view of the HMD significantly distorts field of view and distance perception and most likely making it difficult to drive at higher speed. The authors state that usage specific projections and mapping methods with other geometric properties as possible future improvements. On the positive side, it is noted that users wore the device for more than an hour during several tests without motion sickness or visual fatigue.



Figure 2.25: FlyViz prototype [33]



Figure 2.26: HMD view of FlyViz during driving on a parking lot [33]

FisheyeVision [34] is a similar setup with a HMD but using fisheye cameras and retains the undistorted central field of view but uses non-linear compression to expand the peripheral field of view similar to aspheric door mirrors described in subsection 2.2.1. It thereby supports stereo vision and natural distance perception in the center simultaneously with a compressed field of view of up to 180 degrees in the periphery enabling target detection as tested in a seated, static experiment. The projection method succeeds in smoothly continuing the undistorted center to the compressed periphery but is not optimized for distance or motion perception in the latter and the curved optical flow and inconsistent object size would most likely not be acceptable during driving.

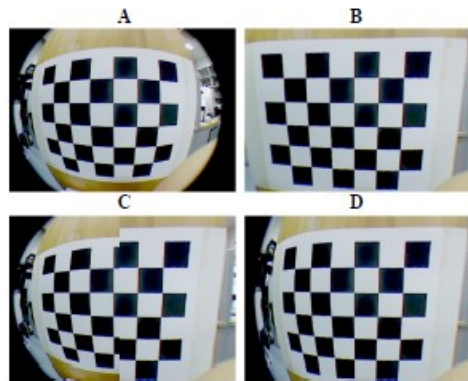


Figure 2.27: The projection used in FisheyeVision (D) compared with the native fisheye camera image (A) and its undistorted counterpart (B) [34]

SpiderVision [35] also uses a HMD and a wide-angle camera but adds an additional camera directed to the rear. Separate conditions were tested for adding rear information into the forward view: semi-transparent overlays of rear objects separated from the background using optical flow, side-by-side images and abstract cues. Compared with the previously described approaches, this one works without compression and therefore without distortion. If used during driving, it may help noticing cars approaching from the rear, but those cars may be displayed onto and occlude the forward view

regardless of the user's intention. In the blended condition preferred by the majority of subjects in their seated experiment, rear objects taken out of context and blended into the forward view may be difficult to locate in space and more confusing for overall situational awareness.

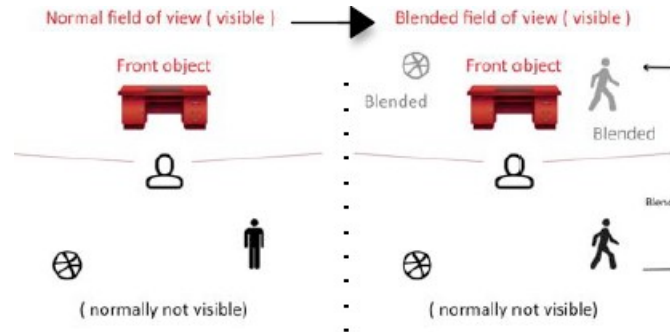


Figure 2.28: Extending field of view by blending [35]



Figure 2.29: Semi-transparent blending (left) and side-by-side display (right) modes of SpiderVision [35]

Outside-In [36] avoids semi-transparency by using picture-in-picture (PIP) previews of off-screen regions-of-interest (ROIs), but PIP can equally occlude the forward view. Yano et al. [37] compared several techniques controlled by head movement under seated and walking conditions and found the effect to be task dependent.

Abstract cues as in the alternative condition in the SpiderVision evaluation may provide sufficient warnings while concentrating on a specific main task and often have the benefit of low workload compared with visually scanning an additional target region. For example, Niforatos [38] proposed augmenting the peripheral perception of skiers' using head mounted sensors and warning lights visible in the periphery in a paradigm that is similar to automotive blind-spot warning systems. Such systems depend on the reliability of sensors, the correct recognition of situations to trigger warnings matching with user expectation and behavior, may still require a final visual check to understand the situation and may help aborting a potentially

---

dangerous maneuver but are often not sufficient for making a positive decision to execute a maneuver.

### 2.3.2 First vs. Third Person Perspective

In some situations, changing the point of view to a third person perspective (3PP) can be preferable over a first-person perspective (1PP). In [39], Gorisse et al. identified 1PP to have a stronger sense of presence, embodiment, sensation of being located in the virtual body, sense of ownership and favored for interactions that require a high degree of precision, while there were no significant differences for sense of agency. In contrast, space awareness and environment perception capacity were higher with 3PP. Further comparisons can be found for example in [40] and [41]. Not every 3PP is equal. [42] explores optimal point of view in 3PP.

3PP is not limited to VR and can be implemented for real life use as in automotive surround view systems described in subsection 2.2.1. LiDARMAN [43] is a mobile 3PP proof-of-concept implementation using a helmet-mounted lidar and HMD that displays a reconstructed view of the surroundings from 1PP or 3PP including a choice of a plan view. The configuration using helmet mounted sensors and a HMD is similar to the FlyViz prototype. While the system is neatly packaged with the computer and batteries in a backpack as a wearable device, with a lidar as the only sensor and the reconstruction neither accumulating 3D data over time nor converting the line scan information into meshes and mapping textures onto them, the surrounding situation is difficult to recognize from the reconstructed view.



Figure 2.30: LiDARMAN [43]

## 2.4 Summary

Human vision in general and while driving is a relatively well researched topic. Research seems to agree on the core mechanisms and important factors.

---

Optimal solutions for visibility during driving, be it something as basic as the size and layout of windows or something more complicated as the size and layout of mirrors, are often trade-offs between contradicting requirements like useful visibility and attractive design, wide field-of-view and large magnification or different situations like slow and fast driving. For forward indirect vision as the main field of view, the consensus seems to be that either a magnification of 1.0 (assuming a wide enough display and field of view) or a well-chosen compromise between large magnification and wide-enough field of view for the given task is optimal for driving performance.

Previous work for automotive systems mostly combined and applied given components like different camera lenses, different mirror curvature or known methods like distortion correction. Attempts to actively design and optimize the perception of space based on ideal requirements that are realized by designing an optimal distortion or selecting an optimal point of view are either non-existent or rare.

Researchers in VR have also attempted solving surround spatial awareness using similar approaches. The proposed methods like field of view expansion and use of 3PP are same as those used in automotive systems. While driving is often mentioned as a possible application, most evaluations were done only indoors in seated or walking conditions in an otherwise small, static environment. Benefit and potential issues for spatial awareness when applied to a task like driving that includes speeds and distances from standstill or parking to highway driving have not been investigated in detail.



---

## 3. Indirect Vision Cockpits

This thesis aims to show that indirect vision has potential to overcome the limitations of natural human perception in a conventional direct vision cockpit and improve spatial perception of the surroundings as envisioned in section 1.3.

While task-independent, wearable indirect vision solutions have a wider applicability, wearable solutions are limited by the weight and size of sensors, computational power and displays. We therefore chose automotive cockpits as our prototype platform as cars are ubiquitous in modern human society and improvements have a large impact on society. While most previous research attempts were mobile but low fidelity or high fidelity but static, cars allow us to use multiple high-resolution sensors and displays that are above what seems to be a minimal level of fidelity for perception under mobile conditions that involve movement and require high spatial awareness. Car driving is a well-defined task with concrete requirements that allows us to develop task specific solutions that can then later be generalized, instead of having to find universal solutions from scratch or improvise unnatural target tasks.

Designs for indirect vision systems can differ in many ways and specifications should include:

- The location, size and field of view of the display within the car and relative to the driver, including whether they are fixed to the interior or to the driver (i.e. HMDs).
- The geometry of the displayed image, described by a projection function that maps three-dimensional space onto the display surface, the point of view (either camera location, the driver's eye point or some other location) and the direction of view. The projection function determines magnification (either smaller, same or larger than life-size; may vary at different locations in the display) and distortion of the image.
- The fidelity of the displayed image, including resolution, frame rate, image delay, color space and color accuracy.
- The presence of binocular and oculomotor depth cues, i.e. stereo and light field displays.
- The abstraction level of the images, for example whether actual camera images or abstract computer graphics are used. The latter could further range from wire meshes to cartoon-like representation. Hybrid solutions

---

could project camera images onto 3D wireframe models that are then rendered from other viewpoints.

- The interaction with the driver, for example if the driver can select the the displayed image from multiple options or whether the direction of view of the image is synchronized to and controlled by body motion.

Which of these decisions are most essential and have the largest impact and should therefore be targeted first?

Some decisions should be postponed because of the lack of technological solutions for the near future. For example, large, high resolution light field displays are not yet available and are not expected for the near future. Other issues like image quality may have a significant effect when improving from below to above a certain threshold for perception, but their continuous improvements make it more of a timing issue than a research topic.

Other decisions may be guided by previous research. For forward view, a life-size magnification of 1.0 seems best above some certain display size at least for parts of the image in which far distance judgments must be made.

Yet other issues like augmentation may be better suited for a separate study not focused on indirect vision systems as they are similarly applicable to direct, see-through displays and difficult to design well.

This thesis focuses on the use and modification of perspective as a unique, powerful and freely designable property of indirect vision systems. It can be designed independent from limitations of currently available technologies like sensor resolution. Concepts and solutions for modifying perspective in automotive indirect vision cockpits may be generalizable, extendable and applicable to other applications.

### 3.1 Projection methods

Projections map three-dimensional space into two dimensions for presenting space on flat displays. Projections determine field of view and magnification and are therefore a core issue of any display system that handles three dimensions. In contrast to the radially symmetric projections of camera lenses, a digital indirect vision system can implement projections that are optimized for human vision without limitations by optical constraints. This includes the possibility of using projections that are not radially symmetric.

Parallel projections are used for example in technical drawings, described by the angle between axes, do not have a particular point of view in space from which the scene is observed and require large display space to cover the visible area needed for driving.

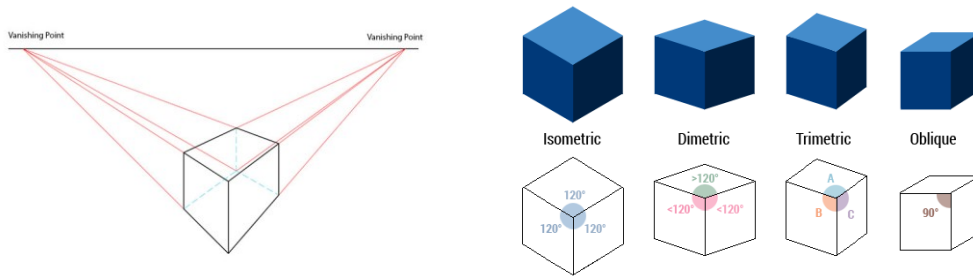


Figure 3.1: Examples for perspective projection (left) and parallel projections (right)

Perspective projections have one or multiple vanishing points and can cover the entire space up to the horizon and the infinitely distant vanishing points on a small display and are therefore suited for displaying the surroundings while driving. They can be described by their point of view in space defined by a location and direction of view and a distortion function. Conventional lenses are radially symmetric, and their distortion can therefore be described by a distortion function that defines the distance from the center of the image on the image plane as a function of the angle from the viewing direction in space.

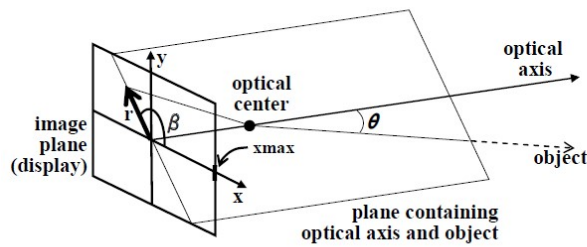


Figure 3.2: The axially symmetric common lens model  $r = F(\theta)$  is independent from the rotational angle  $\beta$ .

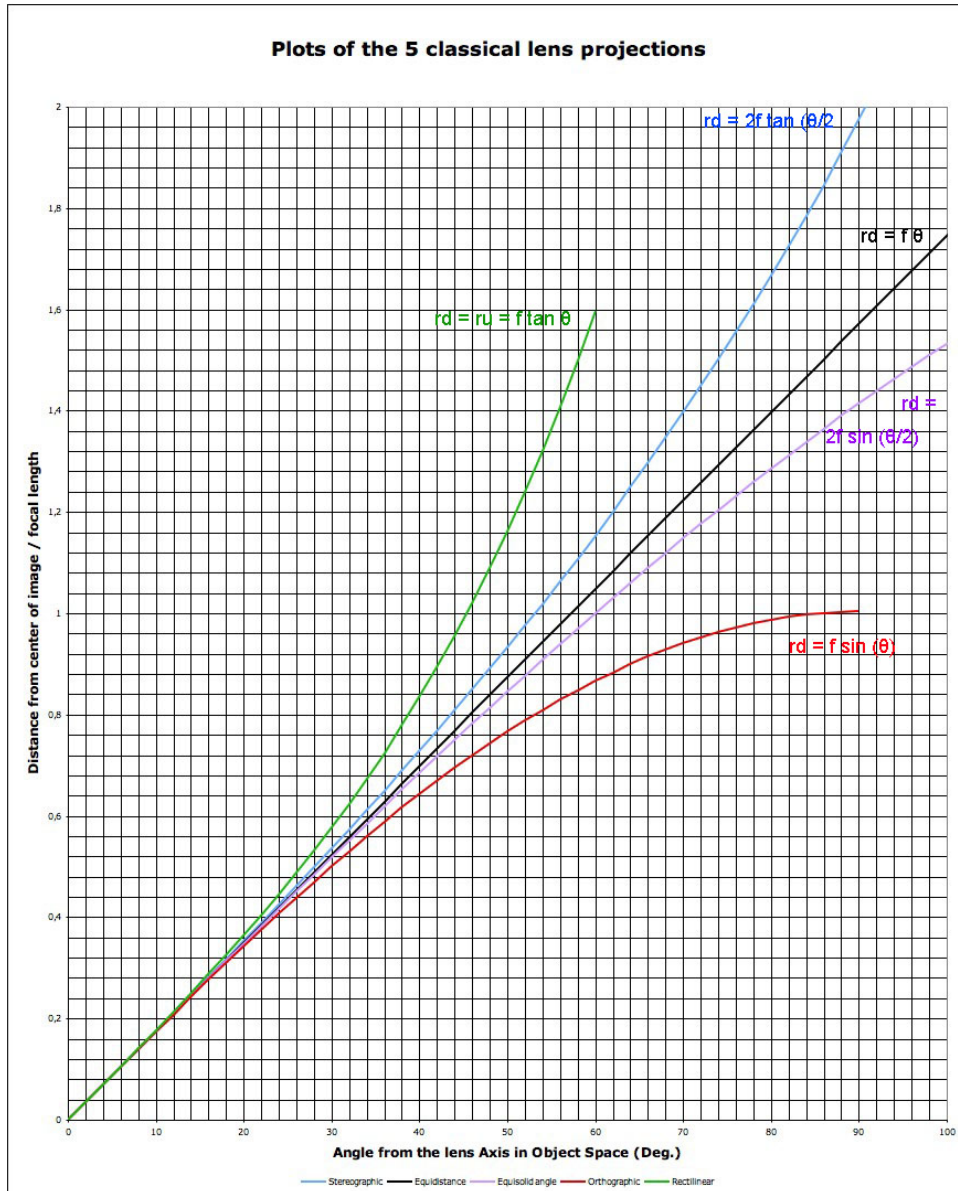


Figure 3.3: Distortion functions of classical lens projections. [43]

Figure 3.3 shows the distortion functions of classical lens projections. Observe that distortion-free rectilinear or pinhole projections which feel most natural for far scenery and narrow field of view require significantly more display size compared with other projections to realize a wide field of view.

### 3.2 Manipulating distortion to expand field of view

Replacing windows by displays gives us the opportunity to intentionally distort the view to provide an effectively larger field of view than geometrically provided by the display. Note that representations of three-dimensional

---

space on a flat two-dimensional display necessarily lack depth causing at least some false oculomotor cues. This can be considered a kind of distortion and aiming for a geometrically distortion-free display should be reconsidered. On the contrary, a stronger distortion can also be an intentional choice as a smaller display with a smaller geometric field of view in the physical space of the user but covering a larger effective field of view may be more efficient to scan for information, especially considering the fact that usable field of view of human perception in which symbols and details can be perceived is only about 60 degrees.

An example are aspherical outside door mirrors which have a higher curvature in the outside area to increase the effective field of view and thereby reduce the blind spot in the next lane. The benefits of this design are that it combines ability to judge far distances at near life-size magnification, combined with a large effective field of view that would otherwise require a larger mirror that is less practical. The outside area is distorted and only allows for judging presence but not exact distance or relative speed, which is an acceptable trade-off for deciding lane changes. Are there projection methods that take the idea of manipulating distortion further utilizing the unique design freedom of electronic indirect vision systems?

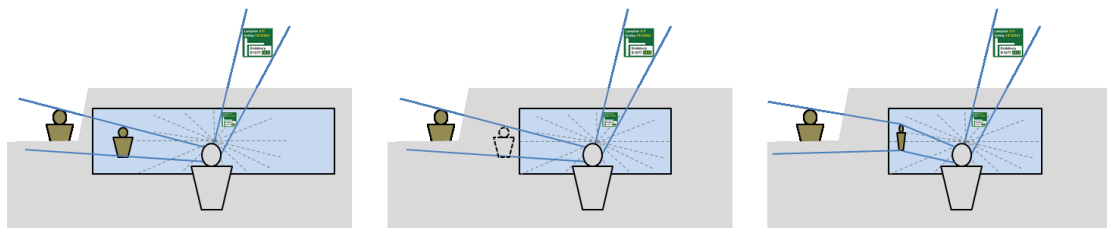


Figure 3.4: Illustration of an intentionally distorted view (right) providing a large effective field of view in a small display. Without such distortion, field of view is a parameter of display size. A large display might not be viewable at one glance. The challenge lies in limiting the unintended side-effects of the intended distortion.

### 3.3 Manipulating point of view

Replacing windows by displays gives us the opportunity to present images from a different viewpoint than the real eye point of the driver. This could be a higher viewpoint showing more context for higher situation awareness or for calming the driver by reduced optical flow, a lower viewpoint showing more detail and providing more thrill from faster optical flow, or a more comfortable backward view when driving backwards for example to park the car.

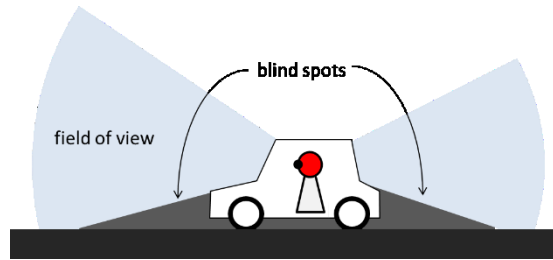


Figure 3.5: Direct vision from the driver's head as viewpoint.



Figure 3.6: Example view from the driver's seat.



Figure 3.7: Illustration of the blindspots around a car.

While human vision and perception handles driving at speeds human beings had never experienced before during the Darwinian evolution process surprisingly well, there are also limitations that make driving difficult. Depending on the given driving situation, the forward view from the driver's eye point might not be optimal. A vision-by-wire cockpit could switch to the optimal viewpoint for the current driving situation to minimize blindspots and provide better overview. In contrast to previous work like Nissan's Around View Monitor system, our approach places the virtual viewpoint image as a natural size main view for driving which we expect to reduce workload.



Figure 3.8: The Vision-by-Wire cockpit showing oneself (the orange car) within the surroundings from a virtual third person viewpoint.

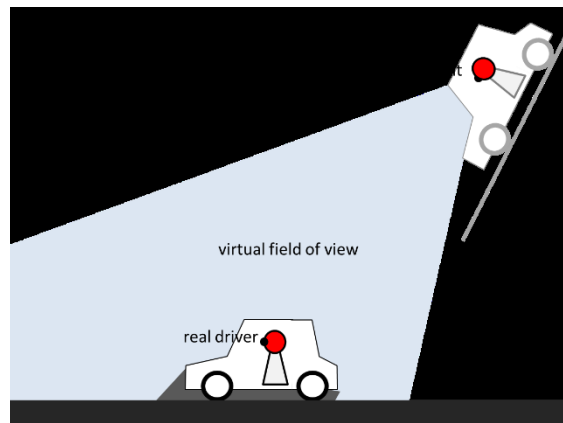


Figure 3.9: Location of real and virtual viewpoints.

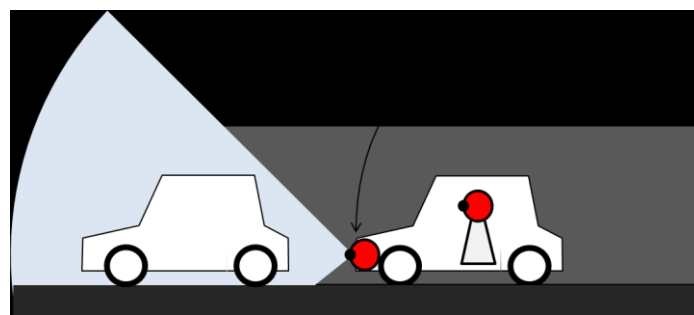


Figure 3.10: Potential of a low viewpoint to reduce close, nearby blind spots.

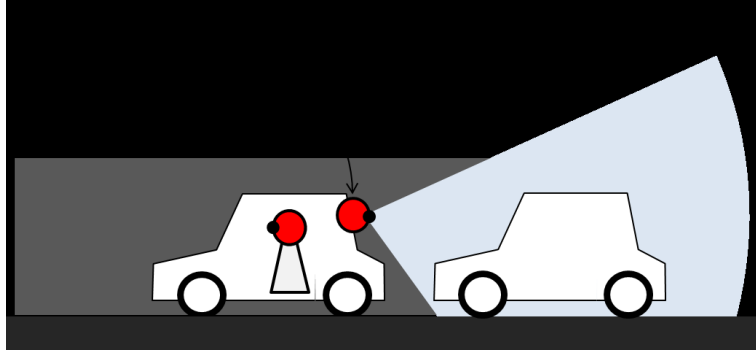


Figure 3.11: A life-size, rearward, virtual viewpoint may be more comfortable for rearward driving than small backup camera displays.

### 3.4 Research plan

The research described in this thesis is structured as follows.

In the first part, comprised of chapters 4 and 5, we first investigate through analysis of the driving task, proposal of new methods, their prototype implementation and experimental evaluation, the possibility to expand the capabilities of natural human spatial perception in conventional cockpits with windows by indirect vision cockpits that replace windows by displays and thereby gain the capability to control perspective.

Spatial perception is needed during driving to perceive and respond timely to the surroundings with suitable actions for safe driving. Perceiving the space along a strip of road centered around the current location towards the direction of travel for navigation and reaching a certain distance backwards to detect faster cars approaching from behind is a basic required task for driving. This requires visual perception of far front and rear simultaneous with a 360-degree view of the near surroundings, which is a much larger than the field of view of about 60 degrees in which humans can recognize symbols at any point in time.

The first focus is therefore on methods to provide a larger effective field of view within the narrow field of view of human perception as sketched in section 3.2. In chapter 4, a new concept for projection is proposed that maps three-dimensional space onto a two-dimensional screen, extending the intuitively perceivable rearward field of view during driving to 180 degrees. This new method was implemented and experimentally verified. It is a significant step from about 30 degrees visible in conventional mirrors. The current implementation is limited to 180 degrees, assumes a straight road section and has been tested for rearward view on a small display, but the method itself is extendable to a wider field of view, to curves and to larger displays.



---

Applying to forward view may require some adaptation as the required properties of the resulting image will partially differ.

Expanding effective field of view is not the only change to perspective that can be achieved in indirect vision cockpits. Another possibility is to modify the point and direction of view to improve surround spatial awareness during driving as described in section 3.3. While it has been claimed that 3PP like plan views improve surround spatial awareness and has been shown to apply to slow speed driving maneuvers like parking using automotive surround vision systems, the potential of 3PP for more general on-road driving situations like curves and intersections has not been addressed in previous work. This thread is followed in chapter 5 by using a HMD-based driving simulator to evaluate and compare the effect of changes in point of view to driving performance and experience.

While chapters 4 and 5 focuses on potential improvements from using indirect vision, using two dimensional displays has potential downsides to spatial perception. In order to reap the benefits, it is essential to identify and offset possible negatives. Previous work, for example, investigated the effect of binocular depth perception, which are small perspective differences between the images captured by the left and the right eye, on driving. Because of the relatively small interpupillary distance, binocular cues work best for near distance. Nonetheless, absence of binocular differences might work as a cue for nearness. When moving our heads, the perspective of both eyes changes continuously. This effect called motion parallax is known to contribute stronger to spatial perception than binocular cues. In order to assess the importance of replicating motion parallax in an indirect vision cockpit, a prototype was implemented and experimentally evaluated.

The research described in the following chapters 4, 5 and 6 therefore provide an overview of the potential of controlling perspective using indirect vision to improve spatial awareness.

---

## 4. Manipulating Distortion<sup>1,2</sup>

### 4.1 Objective

This chapter investigates the possibility of using an indirect vision system to improve spatial awareness compared with direct vision by changing perspective by manipulating distortion through the used projection method.

### 4.2 Example Application

Changing lanes is a maneuver necessary when driving. Driver's need to be aware of cars in the target lane to safely change lanes. The recommended standard procedure is to frequently check far rear traffic in the inner rear mirror which usually covers only about 10 degrees to both sides of the rear for overall rear spatial awareness, before checking the door mirror on the target side which covers only about 20 degrees field of view to verify the empty spot followed by turning around immediately before changing lanes for a final check which is obstructed by the body of the own vehicle. This procedure requires a lot of attention but is still prone to errors. Judging relative speed and there by time-to-contact in the door mirrors is difficult because of the head-on perspective. Recent blind spot warning systems can reduce the need for frequently checking the rear mirror but do not eliminate the need for checking the image in the mirrors followed by turning around.

In this chapter, we investigate the possibility of displaying a field of view of 180 degrees, the maximum recordable field of view by a single camera lens, while still enabling correct distance and speed judgments in the center, detail part of an electronic rear-view mirror. A prototype was tested on roads within our research facility and on public roads.

It has become increasingly difficult to satisfy growing aerodynamic, safety and design requirements simultaneously with a good rear view through large windows divided only by narrow pillars. While modern driving aids

---

<sup>1</sup>T. Yanagi, N. Shimomura, S. Chinomi, H. Mouri, A Novel Camera-Based Rear Vision System for Lane Changes with 180 Degree Field of View, FISITA, 2010.

<sup>2</sup>T. Yanagi, N. Shimomura, An Image Transformation Method that Realizes a Field of View of 180 Degrees without Compromising Distance Perception for a Camera-Based Rear Vision System for Merging and Changing Lanes, First International Symposium on Future Active Safety Technology toward zero-traffic-accident, September 5-9, 2011, Tokyo, JAPAN

like blindspot warning systems and distance sensors warn the driver before getting too close to other traffic and mitigate the restricted rear visibility found in modern cars, such systems generally convey less information than a visual overview. Camera systems that allow displays and small cameras to be placed freely may solve this problem, but how exactly should such a system be designed? How many cameras and displays are required to provide a high level of rear situation awareness at low driver workload?

Solutions for low speed driving like reversing cars into and out of parking lots have become common, but camera systems for merging and changing lanes for a wider speed range from standstill to driving on German Autobahns are less well understood. Show cars often carry rear view systems that simply replace each of the three conventional mirrors by a separate camera-and-display pair. While such systems may satisfy aerodynamic and design requirements, they are costly and do not provide attentional or workload benefits over conventional mirrors: drivers still must look at three separate displays in addition to looking forward and then mentally integrate those separate images into a coherent model of the surrounding situation.

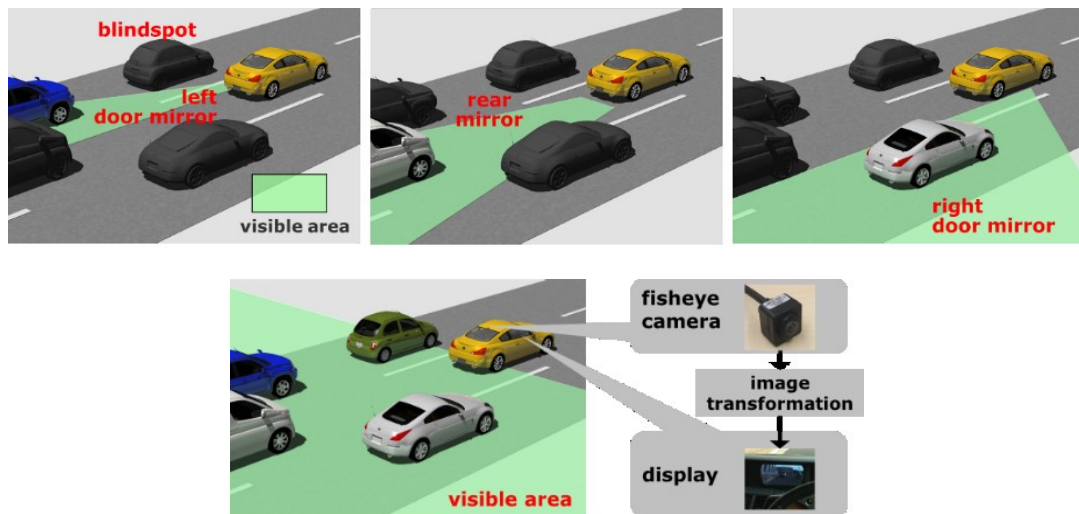


Figure 4.1: Comparison of the field of view of conventional mirrors (top; planar left door and inner rear mirror, convex passenger side door mirror on right) with that of the proposed wide-field-of-view rearview image (bottom).

### 4.3 Method and implementation

We considered different cameras, displays and their combinations with camera view angles ranging from about 40 to over 180 degrees and display sizes from 7 to 11 inches diagonal, and finally settled on a system that uses one fisheye camera mounted in the rear center of the vehicle as the image source.

How should that image be projected onto a display and how large would the display have to be to result in an intuitive representation of up to 180 degrees field of view?

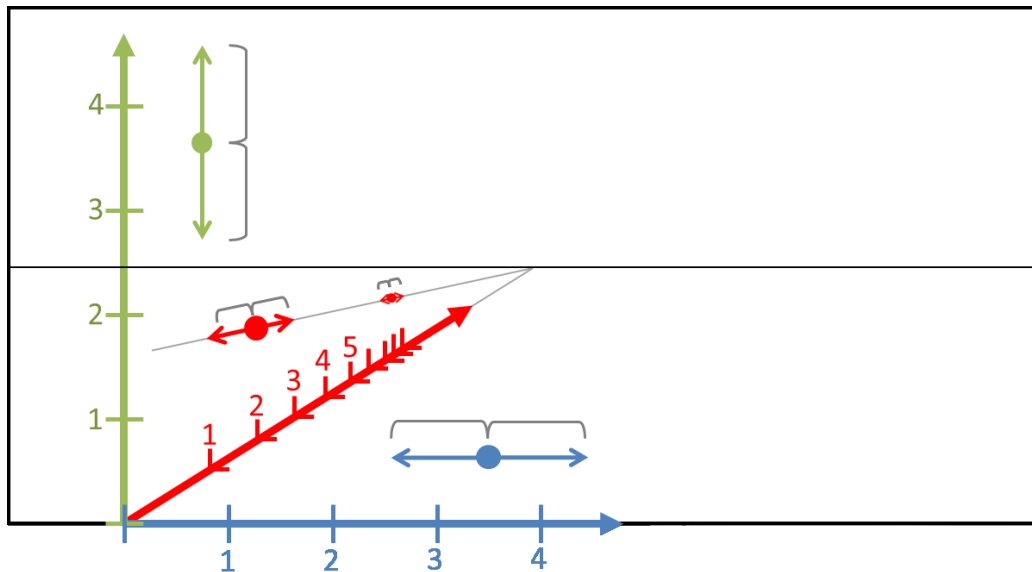


Figure 4.2: Predictability of motion in space when shown on a flat screen display using rectilinear projection.

Consider a rectilinear projection. From the properties of the human visual system, we know that straight line 2D motion at constant speed is predictable. If we consider straight line 3D motion at constant speed, this means that motion in the vertical-lateral plane is predictable, whereas motion in longitudinal direction is somewhat predictable in the far but not in the near, as the rectilinear projection distorts and elongates longitudinal near space.

From this observation, we separate the task into rectilinear projection of the far rear center and a yet to develop projection of the near rear sides. These parts are then joined together similar to aspherical mirrors combining far and near regions.

#### 4.3.1 Far rear center

For the far rear center, rectilinear projection seemed good enough for a first try and the only question was magnification. Magnification is a trade-off between field-of-view, recognizability of small objects and distance perception. Based on previous research, it seemed best to settle on slightly smaller than life-size.

---

### 4.3.2 Near rear sides

A naturally appearing representation of 180 degrees field of view on a flat display has not been attempted before. Keeping in mind any solution is task-specific, we identified the following required properties:

- Objects in the near should be larger than objects farther away, both in longitudinal and lateral direction.
- The location of ground contact should be visible as much as possible also for near objects.
- Nearby objects in neighboring lanes that may possibly be faster than oneself and therefore including bikes should appear large enough to be recognizable.
- The optical flow from self motion should be straight and not distracting.
- The longitudinal distance and relative speed in space from the rear of the own vehicle to nearby objects in neighboring lanes should be easily judgeable.

Fisheye cameras usually satisfy only the first two requirements. Rectilinear projections fail in the last requirement. None of the common projections that were checked satisfied all requirements.

Utilizing the fact that the projection model does not need to be axially symmetric in an electronic indirect vision system, a distortion was designed that compresses the horizontal field of view while preserving straight verticals and straight optical flow during straight ego-motion that satisfies all above requirements.

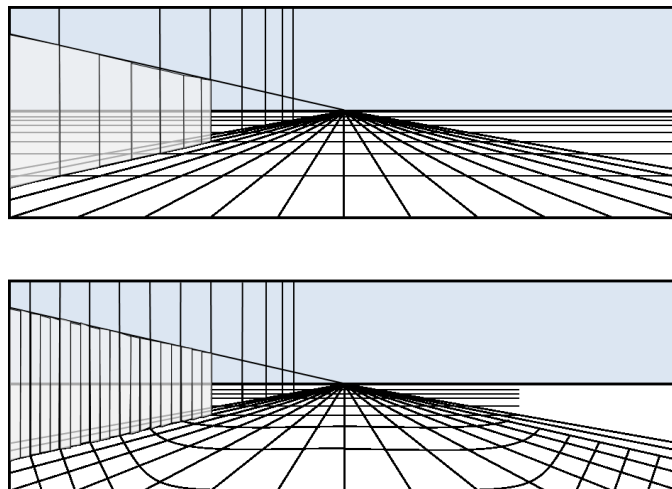


Figure 4.3: Illustration of the idea behind the proposed distorted projection (bottom) compared with a rectilinear projection (top).

Consider a situation in which a car is following 5m behind on the right neighboring lane and closing in to 4m as shown in the following figure. In a conventional right door mirror, that car is about to disappear into the blind spot. Judging relative speed is difficult as the point of ground contact is not visible, the part of the own body visible in the mirror provides no cue about the relative distance. The only usable cue for judging distance and speed is comparing the image to past experience. In a conventional inner rear mirror, this car is already almost no longer visible.

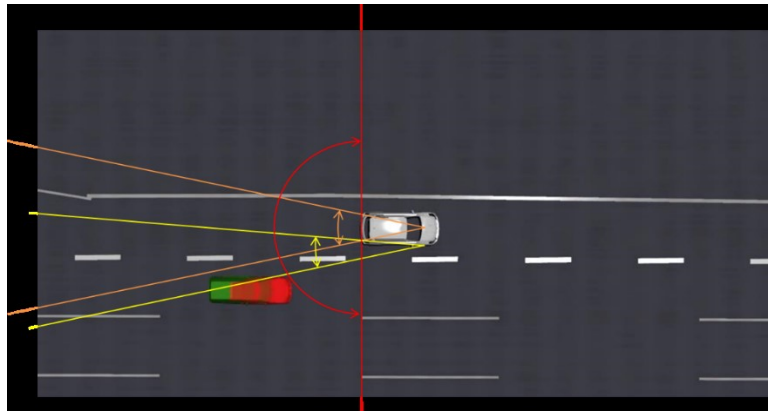


Figure 4.4: Simulated situation with a car following on the right, neighboring lane at 5m (green) and closing to 4m (red). The own vehicle is shown in gray in the center, with lines indicating the field of view of the door mirror (yellow), the inner rear mirror (orange) and the proposed electronic image (red).

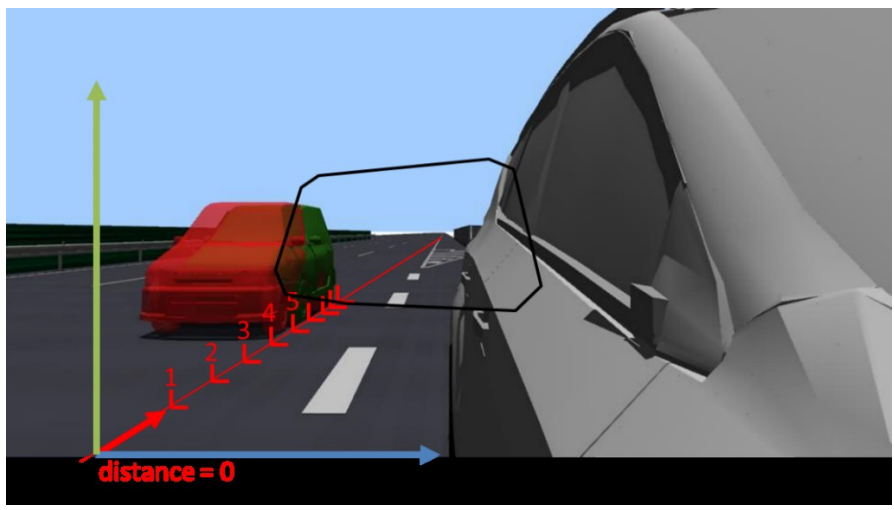


Figure 4.5: Simulated image with about 60-degree field of view with the area visible in a conventional, right door mirror marked in the center.

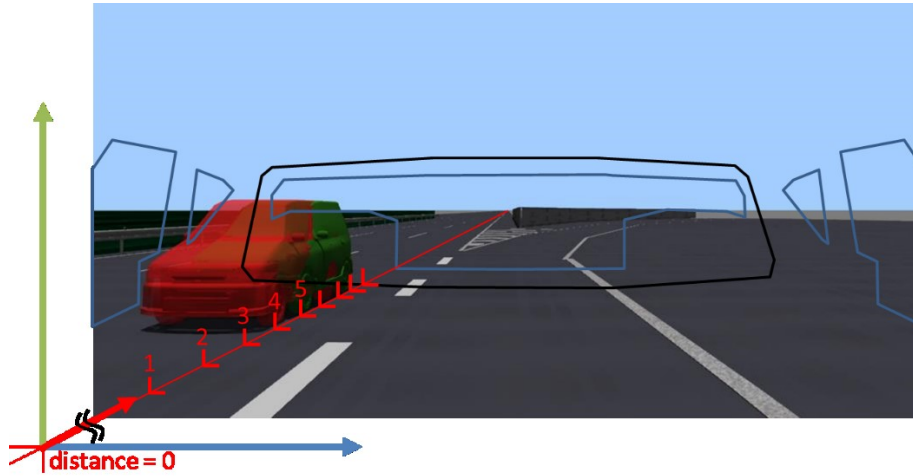


Figure 4.6: Simulated image with about 60-degree field of view with the area visible in a conventional inner rear mirror marked in black and the actually visible area after subtracting the blind spots from the body and rear headrests in blue.

The following figure shows what the proposed image would look like in the same situation. The approaching car is fully visible. Longitudinal motion in 3D space is proportional to horizontal distance from left and right image borders with the borders indicating the location of the rear of the own vehicle, making time-to-contact judgments easy.

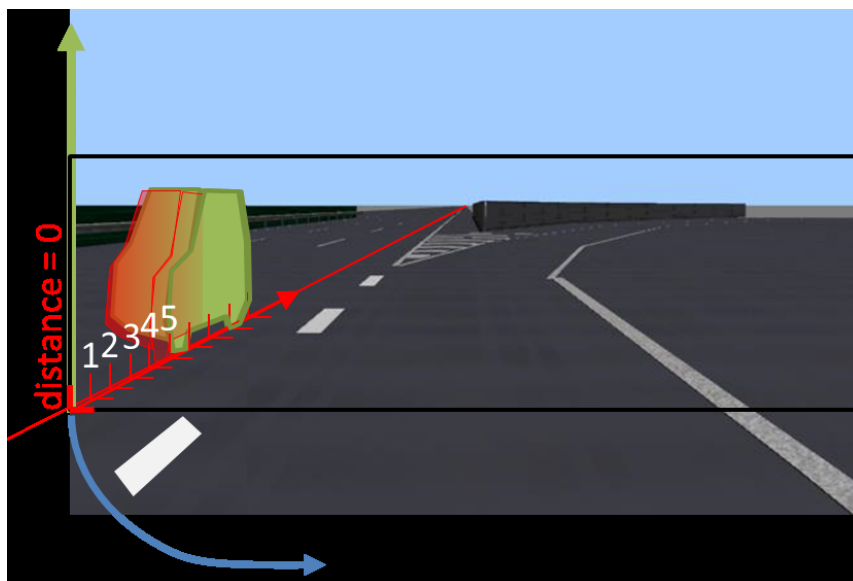


Figure 4.7: Simulated image of the proposed projection with about 180-degree field of view from left to right edge.

Table 4.1 : Comparison of image properties

properties \ projection type		projection type			
		fisheye	perspective	panoramic	proposed
(a)	horizontal field of view of 180 degrees	✓	— <sup>*1</sup>	✓	✓
(b)	(b1) magnification similar to mirrors	—	—	—	✓
	(b2) proportion of far objects	✓	✓	✓	✓
(c)	straight longitudinal lines	—	✓	—	✓
(d)	straight horizon	—	✓	✓	✓
(e)	straight vertical lines	—	✓	✓	✓
(f)	straight transversal lines	—	✓	—	—
(g)	optical flow during straight driving does not change ...	—	✓	✓	✓
	...direction	—	—	—	✓
(i)	proportion of near objects	—	✓	—	—

\*1: Perspective projection is defined only for field of view smaller than 180 degrees. Here we assume one having about 175 degrees.

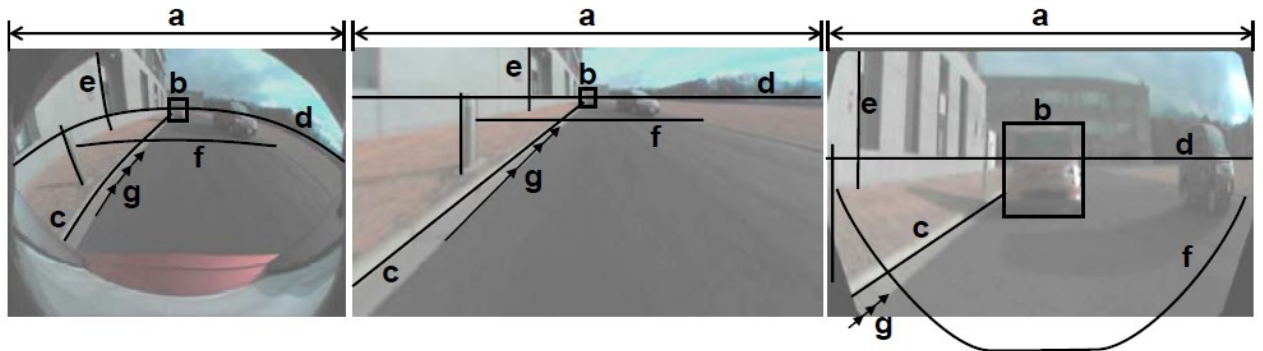


Figure 4.8: Illustration of image properties in the different projections

We extended the rotationally symmetric projection model of a camera described in general by the formula

$$r = F(\theta) \quad (1)$$

where  $\theta$  is the angle between the optical axis and the incoming ray and  $r$  is the distance between the image point and the coordinate origin, to the more general form

$$r = F'(\theta, \beta) \quad (2)$$



that includes virtual cameras that are not rotationally symmetric, where  $\beta$  is the rotation angle around the optical axis.

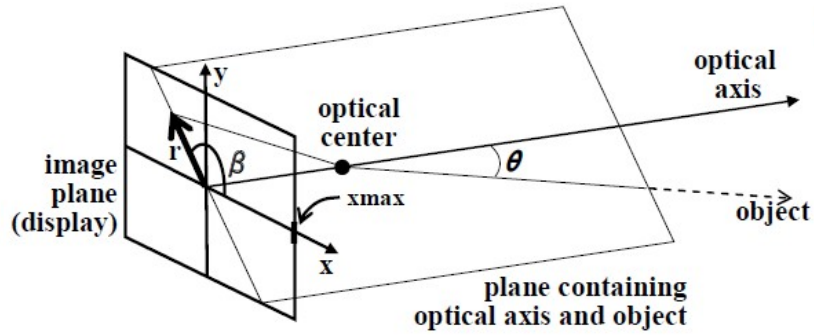


Figure 4.9: Extension of the axially symmetric common lens model  $r = F(\theta)$  to depend on an additional parameter  $\beta$  resulting in a more flexible lens model  $r = F'(\theta, \beta)$ .

Using this representation, properties (c) and (d) are implicitly satisfied by directing the optical axis at the rear vanishing point. The final projection model  $r = F_4(\theta, \beta)$  was numerically approximated starting with a rectilinear projection and applying transformations that implemented the specified requirements. Details of the calculation and the prototype implementation are described in the appendix.

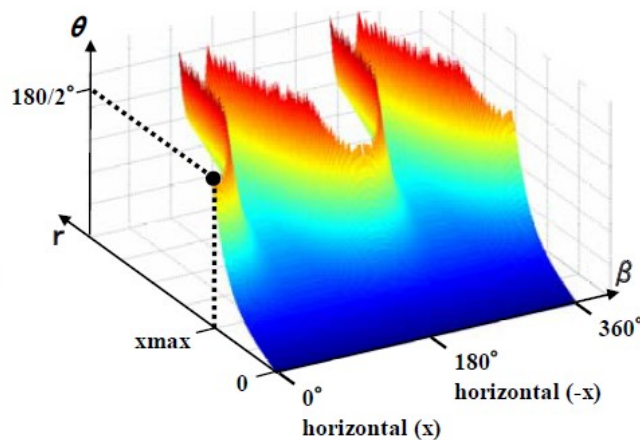


Figure 4.10: The proposed projection  $r = F_4(\theta, \beta)$ .

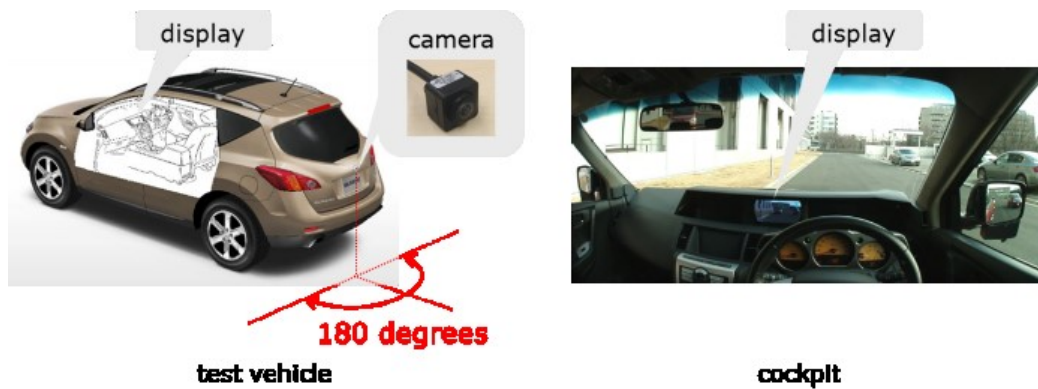


Figure 4.11: Illustration of the complete system. In-camera processing means that only a camera and display are needed.

#### 4.4 Experiments and Results

The following figures show example images from our prototype systems. Manual measurements confirmed that the images satisfy the targeted properties.



Figure 4.12: Sample images before and after processing.



Figure 4.13: A larger sample of the image after processing

---

#### 4.4.1 Distance perception in experiment setup

Distance perception was compared with conventional solutions. The experiment design is based on the assumption that forward distance perception is correct, and that rearward distance judgment should match forward distance perception.

##### Participants and Equipment

Seven male drivers with normal or corrected to normal vision were recruited from our department, varying in age from 29 to 53. Participants sat in the driver's seat of a 2006 Nissan Murano experiment vehicle that was equipped with our prototype system. Four experiment conditions were tested:

- a standard planar rearview mirror,
- the standard convex spherical driver-side door mirror of the experiment vehicle
- a generic rearview camera for reverse maneuvers, and
- our prototype system.

The latter two were displayed on a 9-inch LCD display measuring 198mm by 112mm and mounted centrally on the dashboard with a small offset to the driver. The viewing size measured in viewing angle from the driver's position was similar for driver-side door mirror and our system, both of which were larger than the conventional rearview camera but smaller than the planar rearview mirror. Nissan Serena minivans were used as targets in both front and rear.

##### Procedure

The experiment method described in [44] and [45] asks participants to judge multiples of the distance represented by a forward reference car. Judging multiples of a distance is either basic distance perception or just guessing. In this experiment, we adopted the method to use variable distances for the forward reference car and simplified the task to simply match rearward distance to a probe with the forward distance to the reference.

First, participants were asked in a training phase to watch both mirrors and one of the cameras simultaneously while a car was slowly approaching two times from about 50m in the rear, to get accustomed to the differences in distance perception between the devices. Next, a reference car was placed 30m ahead of the experiment vehicle. Then, a car in the back of the

---

experiment vehicle started to drive forward from an undetermined distance in the rear, until the subject indicated by pressing the brake pedal that the rear vehicle appeared to him to be at the same distance as the car in front. The driver of the rear vehicle then recorded the distance to the experiment vehicle (referred to as judged distance) by reading from a tape measure on the ground. (All distances were measured from the eye-point of the driver.)

This was repeated for six forward reference distances from 30m down to 5m, constituting one block of trials. The two shortest distances were in the near rear side area, whereas the longer distances were in the far rear center area. Four test blocks were completed, one each for the four experiment conditions. The order of the experiment conditions was fixed, starting with the rearview mirror, then the door mirror, then the conventional camera and finally our proposed system as the task was simple and the performance not likely to improve in the later conditions. In each experiment condition, all rearview devices except for the one being tested were hidden by a cover.

## Results

The judged distances for each reference distance in each condition averaged over all participants and the relative error of the judged distances is shown in the following figure. Judged distances for the rearview mirror are closest to correct distance judgements with a relative error of 8%. Using the conventional rearview camera, distances were overestimated by 30%, whereas the driver-side door mirror and our proposed system both resulted in slight underestimations of 23% and 22% respectively. When compared with the conventional rearview camera, the difference to correct distance judgement was smaller with our system for all reference distances except 30m. Standard deviation of the judged distances was smaller for the mirrors than for the cameras and roughly proportional to the reference distance. In the proposed image, there was no significant difference between error in perceived distance in near rear side area and far rear center area, and there was also no apparent effect when transitioning from the far rear center area to the near rear side area. Participants with experience of using conventional rearview cameras were more exact in that condition compared with first-time users.

The results are discussed in chapter 8.



Figure 4.14: View from the driver's seat



Figure 4.15: How the rear was visible to participants.

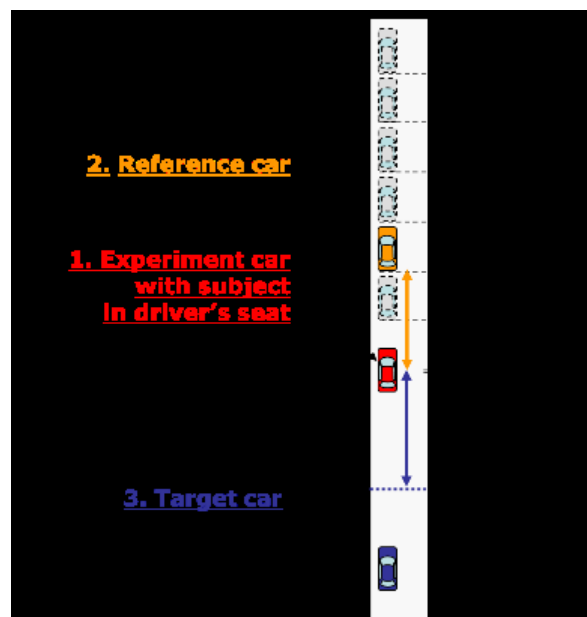


Figure 4.16: Experiment setup.

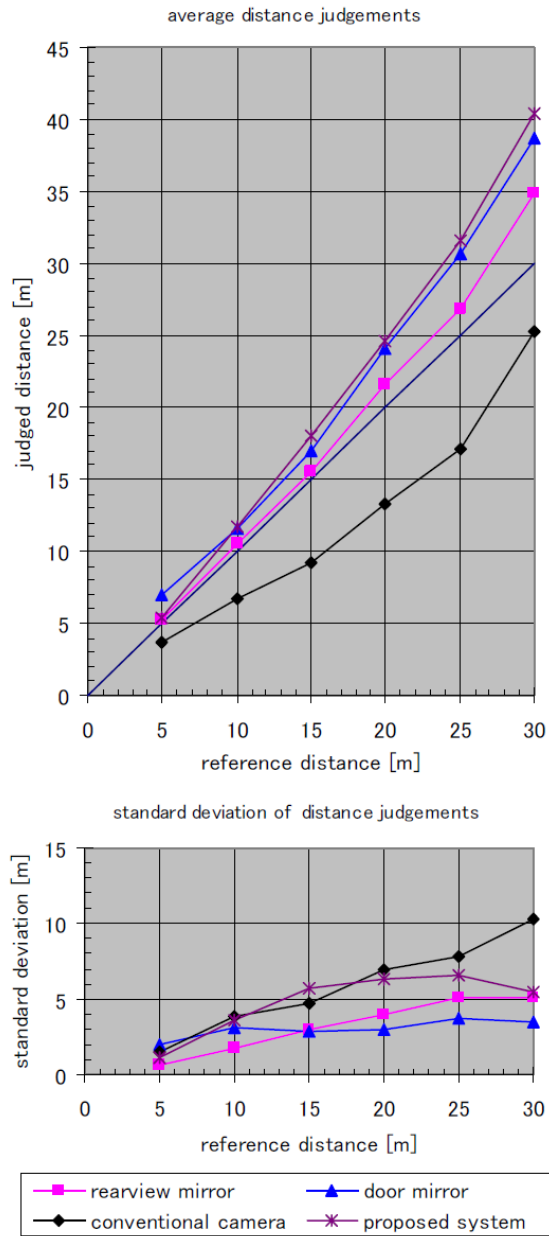


Figure 4.17: Detailed results

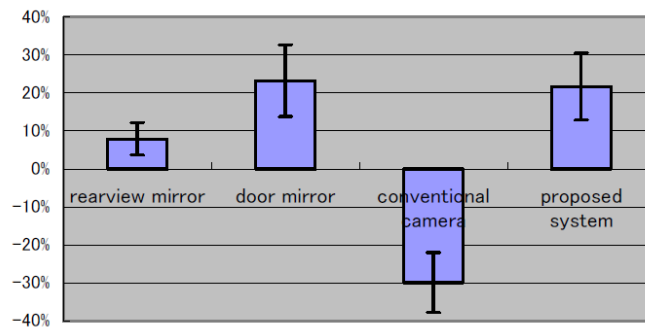


Figure 4.18: Distance judgment error

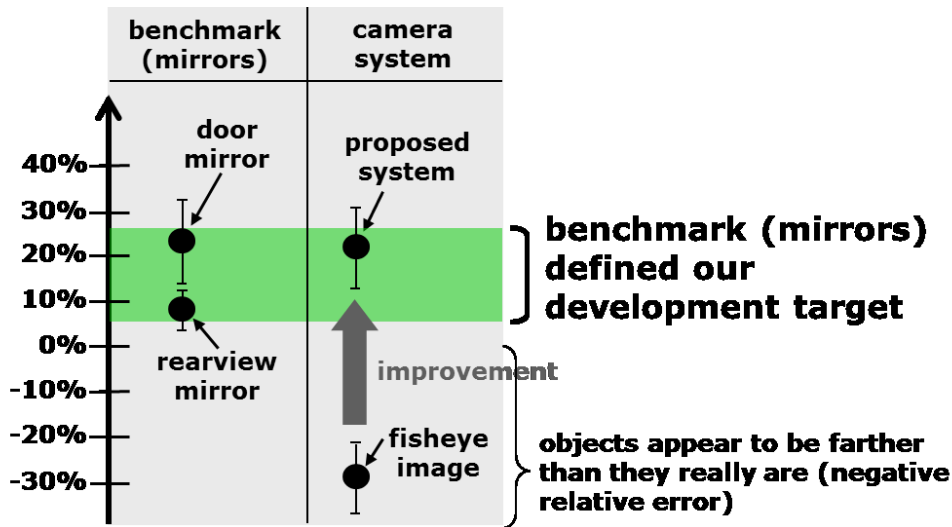


Figure 4.19: Interpretation of the results

#### 4.4.2 Experiments in real driving

Finally, the system was tested during driving on a closed test course and on public roads including lane change and merging maneuvers. Participants were positive about the system and reported that the complete view of the rear in one location as presented by this system helped them obtain awareness of the complete surroundings by looking ahead and at the display, without large head and eye movements required by mirrors. The natural appearance of the image achieved by correcting unwanted distortions was also rated positively. Images captured with our prototype during a lane change on public roads are shown in the following figure and illustrate how the complete rear situation is visible during the whole lane change. The naturalness of the optical flow after transformation when driving straight forward was also confirmed during driving.

On the negative side, the image seemed unnatural for ego-motion other than straight longitudinal motion including rotational motion in curves and pitching when driving over bumps. Some participants missed some orientational guidance that helped them identify the direction of the image or to judge distances relative to some threshold.



Figure 4.20: Sample images of the prototype from driving on public roads.

---

# 5. Manipulating Point of View

## 5.1 Objective

The previous chapter investigated the possibility of actively using distortion to improve spatial awareness. This chapter focuses on actively choosing point of view to improve spatial awareness in an attempt to answer the following questions:

- Can varying the point of view improve driving performance and experience by improving spatial awareness?
- Can the point of view be manipulated with current technologies?

Smaller changes in viewpoint according to the motion of the user's head in order to replicate motion parallax will be discussed in the next chapter.

## 5.2 Which Viewpoints?

Out of the possible viewpoints including first-person and third-person viewpoints as well as forward, rearward and top-down viewpoints shown in the next figure, only a subset provides seeing the own car and the direction of travel which is necessary when used as the only view. Those viewpoints roughly line-up on the red line shown in the second figure.

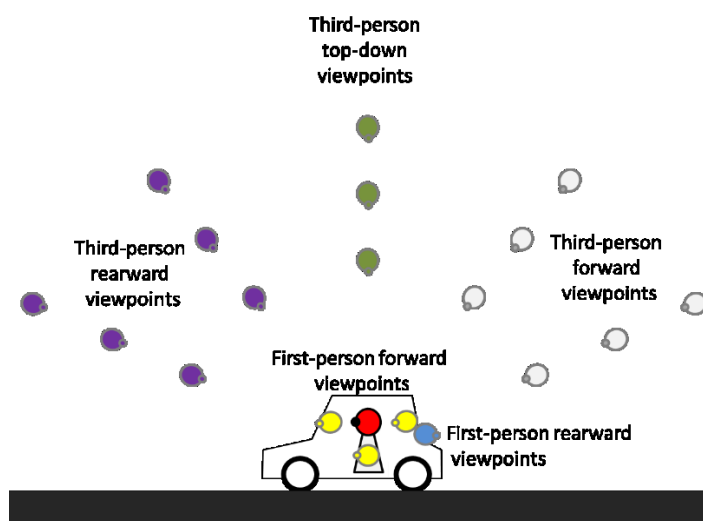


Figure 5.1: Possible viewpoints.



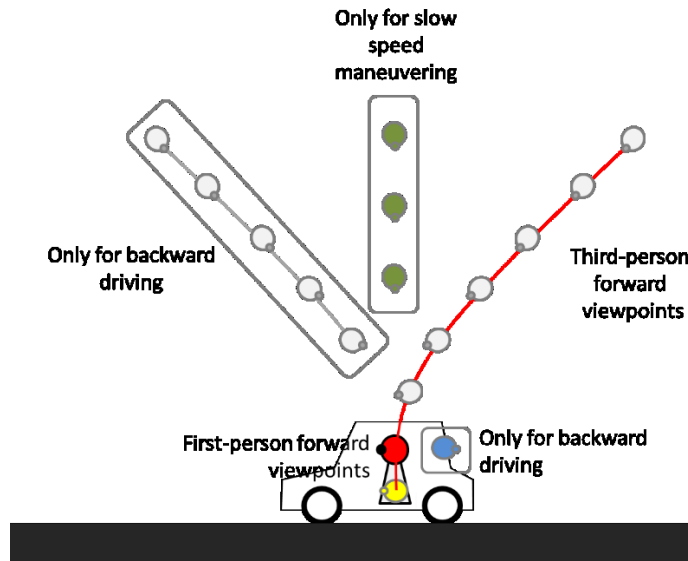


Figure 5.2: Viewpoints that can be used during forward driving as the only viewpoint.

Additional viewpoints are possible shifted laterally to the left or right and may be useful for example to see beyond curves and corners but are left as future work as they can be added as extensions later on.

### 5.3 Simulator experiment

A simulator experiment was completed to answer whether varying the point of view can improve driving performance and experience by improving spatial awareness.

#### 5.3.1 Experiment design

In addition to the normal driver viewpoint, a higher birdview-like viewpoint for surround situational overview and a lower racing car-like viewpoint for better visibility of low obstacles were implemented as the two extremes of the viewpoints described in Figure 5.2. While the normal and low viewpoints were experienced from within the cockpit without changing the eyepoint relative to the own car, the high viewpoint was implemented without the own car surrounding the user viewpoint because of the limited resolution of the HMD used. One of these viewpoints was active at each moment and selected by user-controlled head pose in the practice session (i.e. looking downwards shifted viewpoint upwards whereas looking slightly upwards shifted the viewpoint downwards) or preset as experiment condition. In a real-world implementation, it may be selected by the user or automatically depending on driving situation. A virtual rear view that provided a backward view to the rear from the rear seats without actually having to turn the head backwards was made available in the low viewpoint condition.

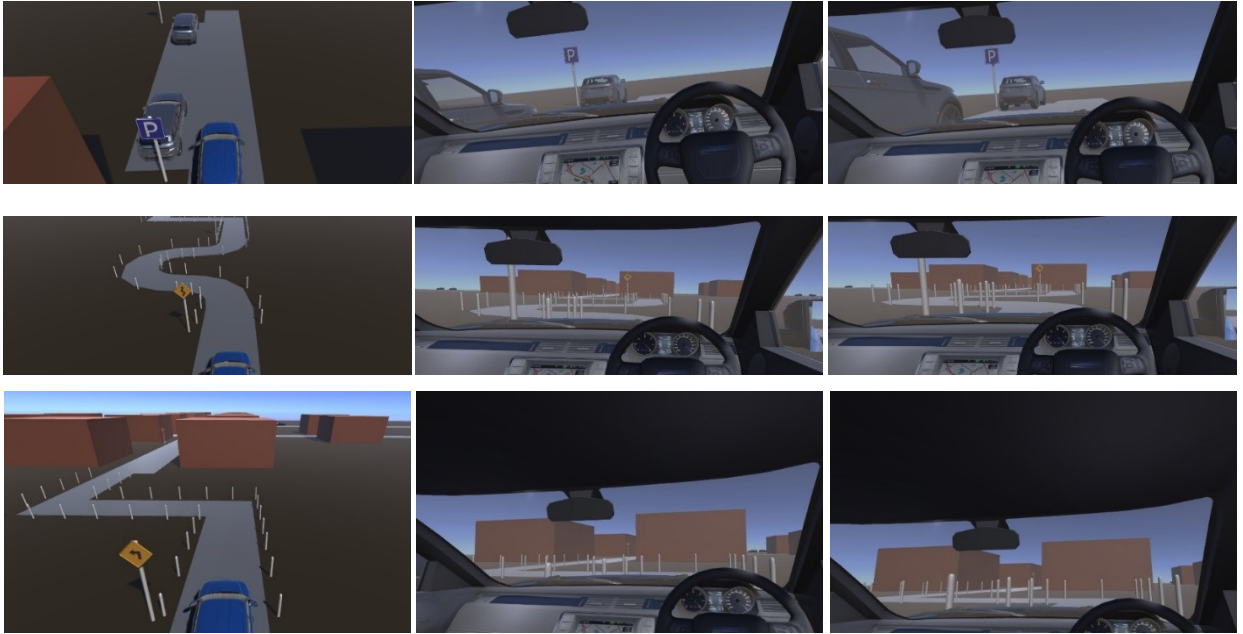


Figure 5.3: Examples of views from the high viewpoint, normal eye point and the low viewpoint.

As an alternative method to improve spatial awareness, we investigated the potential of using Augmented Reality. Many possibilities exist for utilizing augmented reality to improve spatial awareness, decision making and control performance. Unluckily, there is no systematic research on best augmented reality practices for driving assistance. As our aim was not a deep investigation of the possibilities, limitations and other details of using AR for driving, we decided to implement just two types of augmentation that we considered essential for situation awareness, which were virtual walls to clearly visibly block roads that one must not enter (e.g. one-way road exits) to demonstrate aiding awareness of static road environment, and virtual bars growing from the front of other cars to indicate safe distance to demonstrate aiding awareness of dynamic surroundings. AR to help controlling the vehicle was not implemented as predicted path trajectories depending on steering angle is already known from present rearview camera systems but also present an unsolved challenge when used for curve driving where a simple implementation of a trajectory depending on current steering angle would quickly go off-road and be more confusing than useful.

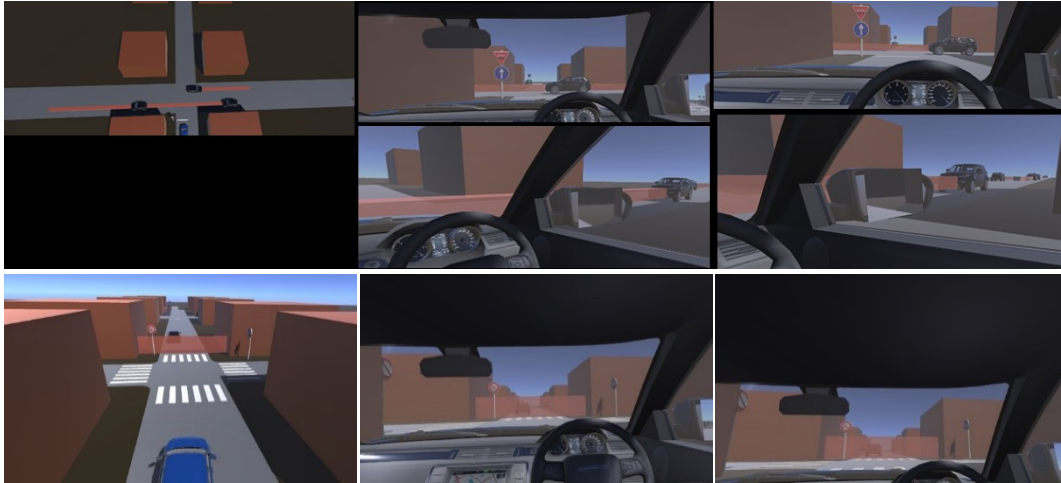


Figure 5.4: Examples of views at an intersection with AR. From left to right, high viewpoint, normal viewpoint and low viewpoint.

### 5.3.2 Experiment Task

Driving encompasses many tasks including road navigation, lane navigation, understanding of static and dynamic (e.g. traffic signals) driving rules, avoidance of static and dynamic obstacles, communication and negotiation with other cars, and control of vehicle motion. We designed a driving course that starts with exiting a parallel parking space where participants were asked to leave without hitting the cars in front and in the back, continues with a tight curve and a narrow bend which were marked by poles on both sides of the road and where participants were asked to drive smoothly without leaving the road, followed by a slow speed “static” intersection with traffic signs either denoting that a road can be entered or not, pedestrians and parked cars where only one out of three directions was allowed to be entered. In the AR condition, one of the two forbidden directions was blocked by a virtual wall reducing the number of options from 3 to 2 – in pre-experiments, this seemed a good balance between too easy (i.e. reducing the number of options to 1) and too difficult, and also appeared to be easy to understand even though this would be a condition of incomplete augmentation in which not all similar cases were augmented. The task continued with another “dynamic” intersection with fast cross-traffic where participants were asked to safely cross the road, and a final parallel parking maneuver. For the dynamic intersection, the AR condition consisted of red, safe-distance indicators in front of the other cars. These were implemented as complete and correct augmentation as our focus was not trust issues but effect of AR on situation awareness and adding those factors would have over complicated this experiment.

The course thereby included driving tasks of all three situation awareness levels and tasks that require up to 180 degrees (e.g. watching cross traffic

from left and right at an intersection) and surround 360 degrees of situation awareness (e.g. when leaving or entering a parallel parking spot). The traffic signs, pedestrians and parked cars in the static intersections were randomized but equal between participants. The cross traffic in the dynamic intersections moved with a prefixed distance pattern between vehicles but the start position of the pattern was randomized in each trial and the length of the patterns from left and right was different, balancing the need for unpredictable patterns with similar difficulty in each trial.

Table 5.1 : Overview of driving tasks in the experiment

	Experiment task	Situation Awareness				Decision-Making and Control		
		Direction	Level 1 Perception	Level 2 Comprehension	Level3 Projection	Knowledge-based behavior (Decision making by prediction and planning based on symbols)	Rule-based behavior (Decision-making by patterns/conditions based on signs)	Skill-based behavior (Automated, sensitive patterns based on signs)
1st situation	Parallel parking exit	<ul style="list-style-type: none"> <li>Near front and rear</li> </ul>	<ul style="list-style-type: none"> <li>Distance towards front and rear obstacle</li> <li>Position/pose relative to opening</li> </ul>	<ul style="list-style-type: none"> <li>Required maneuver (steering &amp; longitudinal movement)</li> </ul>	<ul style="list-style-type: none"> <li>Ideal and expected Trajectory</li> <li>Interaction with surrounding traffic</li> </ul>	<ul style="list-style-type: none"> <li>Decide how to deal with surrounding traffic</li> </ul>	<ul style="list-style-type: none"> <li>Standard maneuver for exiting parking slot</li> </ul>	<ul style="list-style-type: none"> <li>Stop close to but before hitting obstacle</li> </ul>
2nd / 3rd situation	Curve / Crank	<ul style="list-style-type: none"> <li>Forward</li> <li>Near left and right</li> </ul>	<ul style="list-style-type: none"> <li>Position/pose relative to road boundary and obstacles</li> </ul>	<ul style="list-style-type: none"> <li>Road curvature and current Trajectory</li> </ul>	<ul style="list-style-type: none"> <li>Ideal and expected Trajectory</li> </ul>	<ul style="list-style-type: none"> <li>Strategy if failed and completely off-road</li> </ul>	<ul style="list-style-type: none"> <li>Trajectory that considers vehicle motion limitations</li> </ul>	<ul style="list-style-type: none"> <li>Speed/steering control</li> </ul>
4th situation	Slow Intersection, decide direction from 3 (w/o AR) or 2 (w/ AR) options	<ul style="list-style-type: none"> <li>Forward, around corner</li> </ul>	<ul style="list-style-type: none"> <li>Traffic signs</li> <li>Pedestrians</li> <li>Cars</li> <li>Positions/pose in intersection</li> </ul>	<ul style="list-style-type: none"> <li>X-type intersection with pedestrian crossings</li> </ul>	<ul style="list-style-type: none"> <li>Ideal and expected Trajectory</li> </ul>	<ul style="list-style-type: none"> <li>Direction allowed</li> </ul>	<ul style="list-style-type: none"> <li>Intersection passing trajectory</li> </ul>	<ul style="list-style-type: none"> <li>Exact stop, turning</li> </ul>
5th situation	Intersection with fast cross traffic (AR: THW decision aid)	<ul style="list-style-type: none"> <li>Forward, around corner</li> </ul>	<ul style="list-style-type: none"> <li>Location and speed of cars</li> <li>Position/pose in intersection</li> </ul>	<ul style="list-style-type: none"> <li>High priority Cross-traffic situation</li> </ul>	<ul style="list-style-type: none"> <li>Anticipate and find gaps in traffic</li> <li>Distance to traffic when crossing</li> </ul>	<ul style="list-style-type: none"> <li>Strategy if no gaps</li> </ul>	<ul style="list-style-type: none"> <li>Go decision</li> </ul>	<ul style="list-style-type: none"> <li>Exact stop</li> </ul>
6th situation	Parallel parking	<ul style="list-style-type: none"> <li>Forward, rearward</li> <li>Near left, front and rear</li> </ul>	<ul style="list-style-type: none"> <li>Distance towards front and rear obstacle</li> <li>Position/pose relative to opening and road boundary</li> </ul>	<ul style="list-style-type: none"> <li>Required maneuver (steering &amp; longitudinal movement)</li> </ul>	<ul style="list-style-type: none"> <li>Ideal and expected Trajectory</li> </ul>	<ul style="list-style-type: none"> <li>Strategy if failed</li> </ul>	<ul style="list-style-type: none"> <li>Standard maneuver for parallel parking</li> </ul>	<ul style="list-style-type: none"> <li>Speed/steering control</li> </ul>

Situation awareness while driving has many different aspects which include understanding and awareness of the route on a road-level and a lane-level, understanding of static (e.g. traffic signs) and dynamic (e.g. traffic signals) driving rules, awareness and avoidance of static (e.g. curb stones) and dynamic (e.g. deer) obstacles, communication and negotiation with other cars and control of vehicle motion. The following table orders driving tasks by

situation awareness level (according to Endsley) and field of view. Experiments should include representative driving tasks from all situation awareness levels as well as different field of view in order to obtain results representative for natural driving.

Table 5.2 : Typical driving tasks

Situation Awareness level \ Field of view	One-direction (up to 90 deg)	(up to 180 deg)	Surround (up to 360 deg)
Level 1 Perception (and Control Adjustments)	Driving, stopping, curves, obstacle avoidance	Merging	Surround situation during parallel parking maneuver
Level 2 Comprehension	Route, obstacles, intention of other cars	Path through intersection	Parking maneuver
Level 3 Projection	Timing of traffic signal, anticipation of movement of other cars	Intersection timing	

Table 5.3 : Overview of the tasks in the experiment

Situation Awareness level \ Field of view	One-direction (up to 90 deg)	(up to 180 deg)	Surround (up to 360 deg)
Level 1 Perception (and Control Adjustments)	Acceleration, braking, stopping, steering	Driving through crank and turning at intersection	Parking maneuver
Level 2 Comprehension	Existence and intention of obstacles in front	Existence and intention of obstacles to the sides, path through intersection	Parking maneuver
Level 3 Projection	Anticipating future situation caused by own movement	Avoiding cross-traffic	Anticipating result of parking maneuver

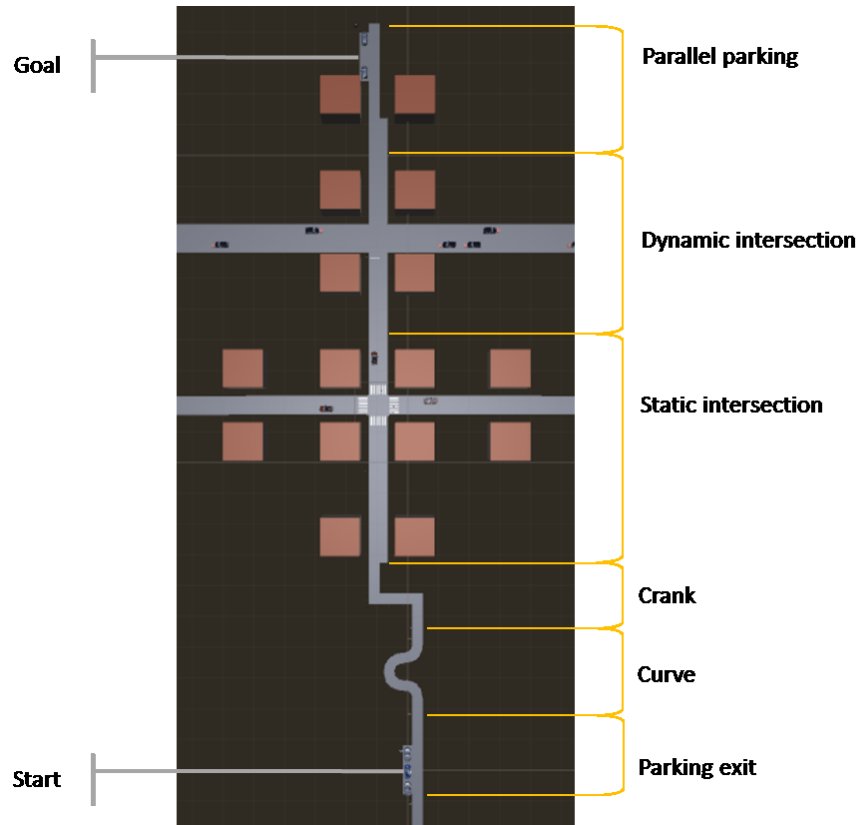


Figure 5.5: Driving course used in the experiment.

### 5.3.3 Results

Figure 5.6 shows the duration spent in each section averaged over all participants. High viewpoint seems shorter than normal viewpoint in dynamic intersection and parking, slightly shorter in parking exit and curve, while slightly longer for crank and static intersection. While the average duration for high viewpoint is at the bottom of the variation of the normal viewpoint for dynamic intersection, for other sections this is not the case. We observe less variation between subjects for high viewpoint.

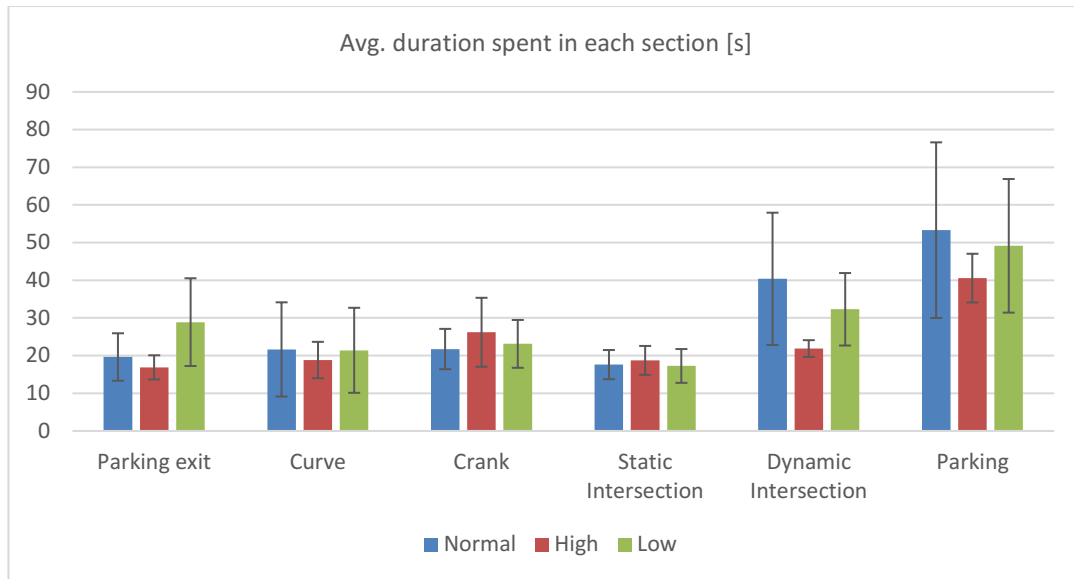


Figure 5.6: Average duration spent in each section.

The overall enquete results show a general preference for high viewpoint in all asked aspects – situation awareness, correctness of decisions, anxiety, fun to drive, mental workload and physical workload.

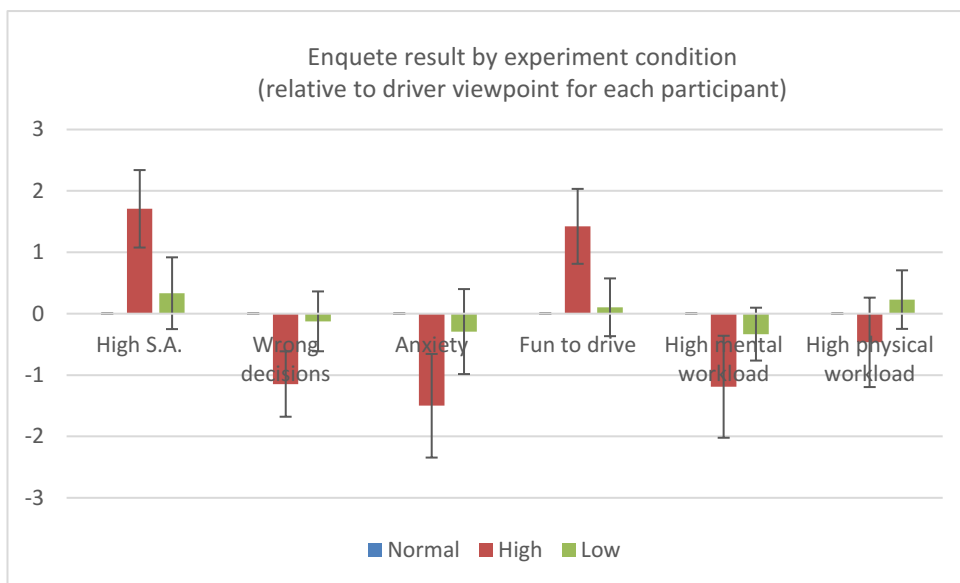


Figure 5.7: Average enquete results relative to normal viewpoint on a per participant basis for all sections

### 5.3.3.1 Dynamic intersection

Figure 5.8 and Figure 5.9 confirm that on a per subject basis, duration for dynamic intersection is shorter for high viewpoint compared with normal viewpoint whereas the duration for low viewpoint compared with normal viewpoint varies between participants.

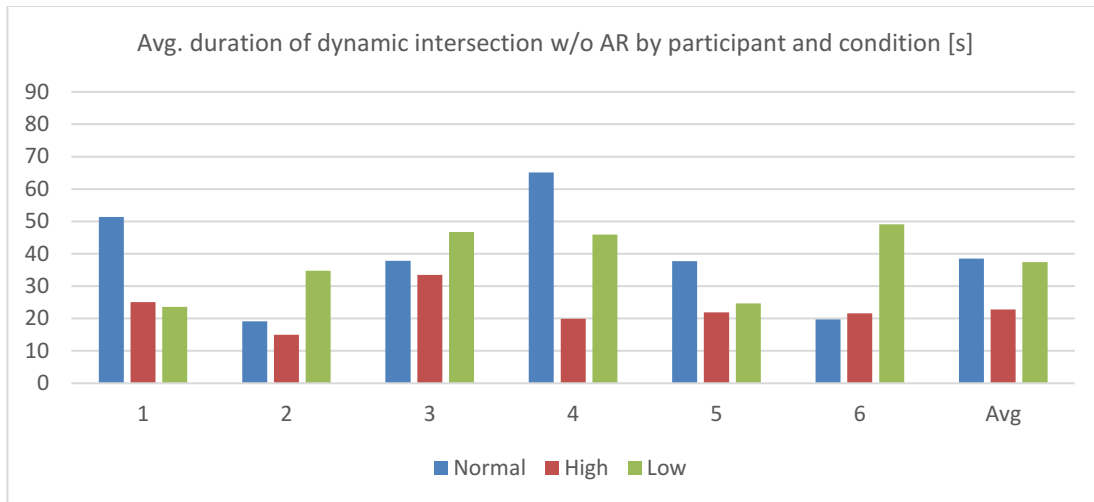


Figure 5.8: Average duration of dynamic intersection without AR.

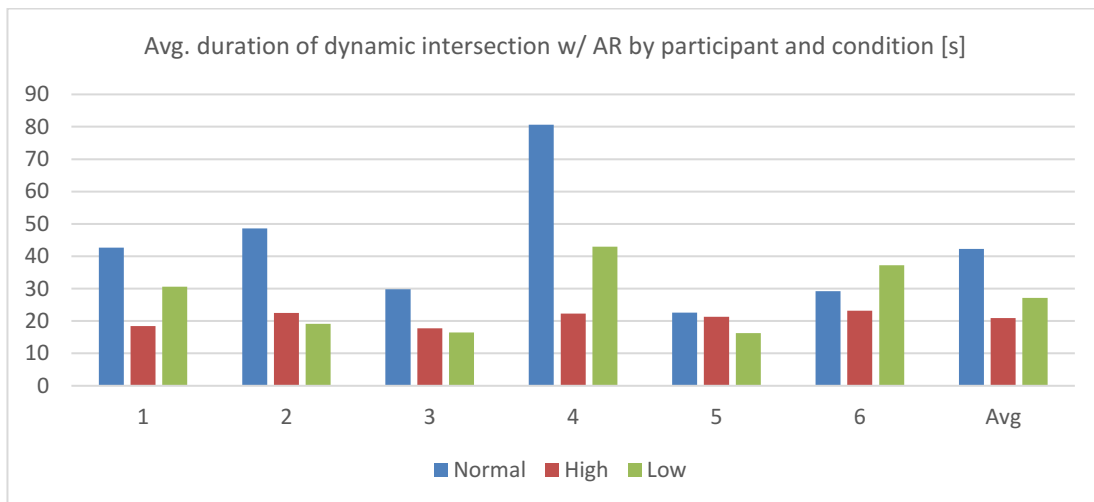


Figure 5.9: Average duration of dynamic intersection with AR.

The total duration of missed opportunities shows a trend similar to duration spent at the dynamic intersection as it is the foremost factor deciding duration spent. A look at the duration of missed opportunities per participant largely shows the same trends. (Figure C.15)



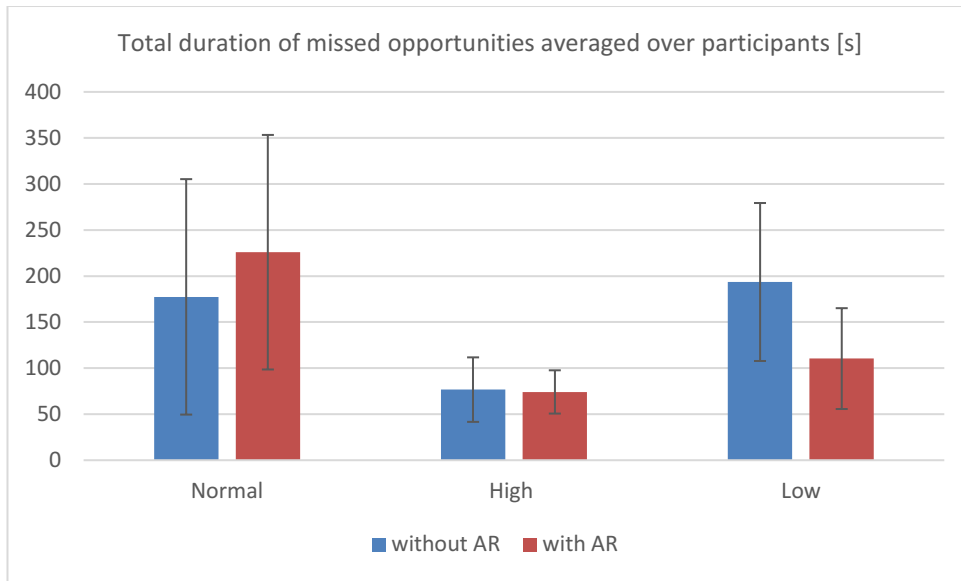


Figure 5.10: Total duration of missed opportunities.

Table 5.23 shows the number of crashes and near misses by participant, further divided into with and without AR. Half of the participants experienced at least one crash. There are more near-misses with AR than without AR.

Table 5.4 : Number of crashes and near misses.

w/o AR - w/ AR		Participant					
		1	2	3	4	5	6
crash	normal	0-0	0-0	0-0	2-1	0-0	0-0
	high	0-0	0-0	0-0	0-0	0-1	0-0
	low	1-0	0-0	0-0	0-0	0-0	0-0
near miss	normal	0-0	0-0	0-0	0-1	1-1	1-0
	high	0-1	0-0	0-0	1-1	0-0	0-0
	low	0-0	0-1	0-0	0-1	0-1	0-0

Figure 5.11 shows the minimum distance from cross traffic to the participant when crossing the cross road. It is similar between no-AR and AR for normal and low viewpoints. For high viewpoints, minimum distance seems shorter in the AR condition compared with no AR.

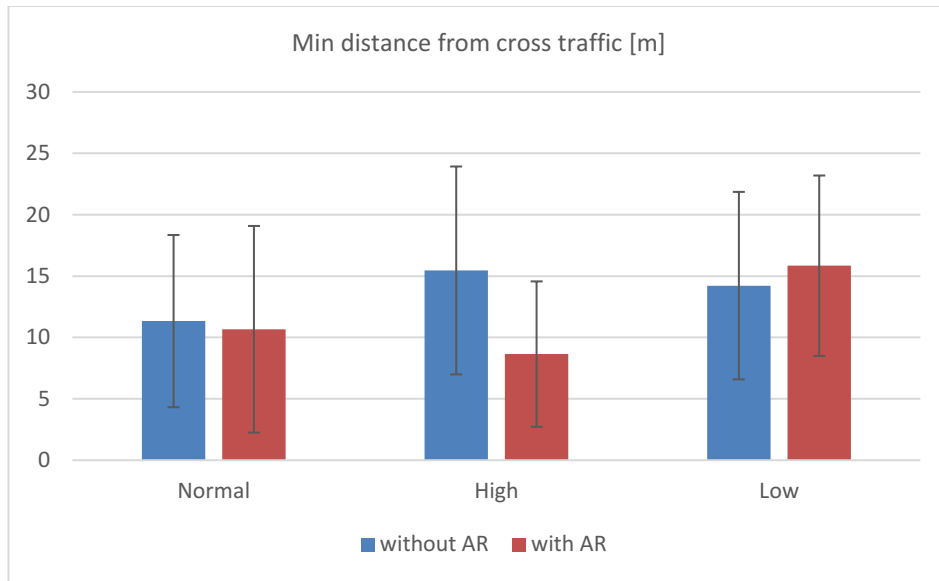


Figure 5.11: Minimum distance from cross traffic to participant when crossing the dynamic intersection, averaged over participants and repetitions.

The range of head rotation around the vertical axes decreases for the high viewpoint by almost half and often slightly also for the low viewpoint. AR does not seem to affect motion range for the static intersection but does so for the dynamic intersection. These trends are observable also on a per participant basis. (Figure C.16)

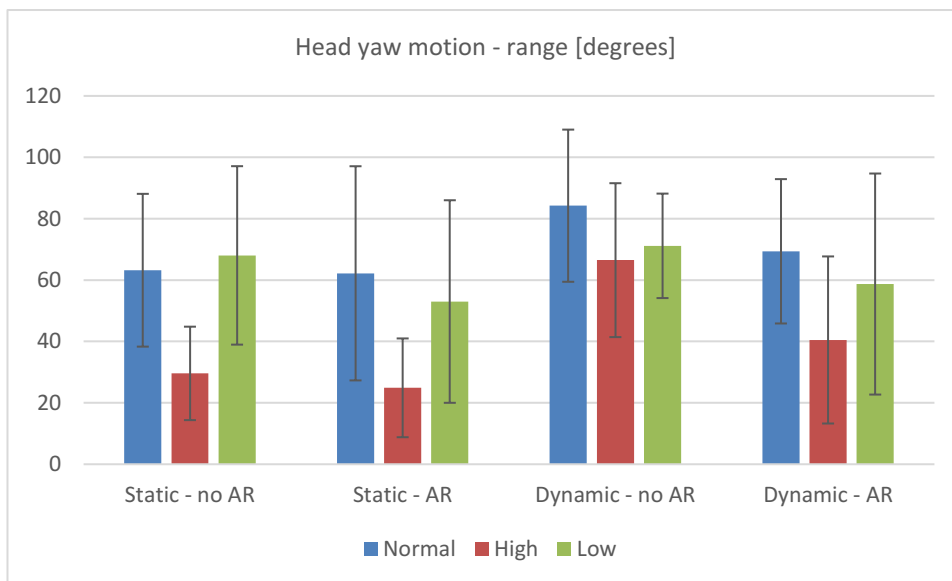


Figure 5.12: Range of head motion around vertical axis averaged over participants.

Subjective reports show preference for high viewpoint in all aspects but almost no difference between AR and no-AR. If anything, the self-rating for correct decisions decreased with AR.

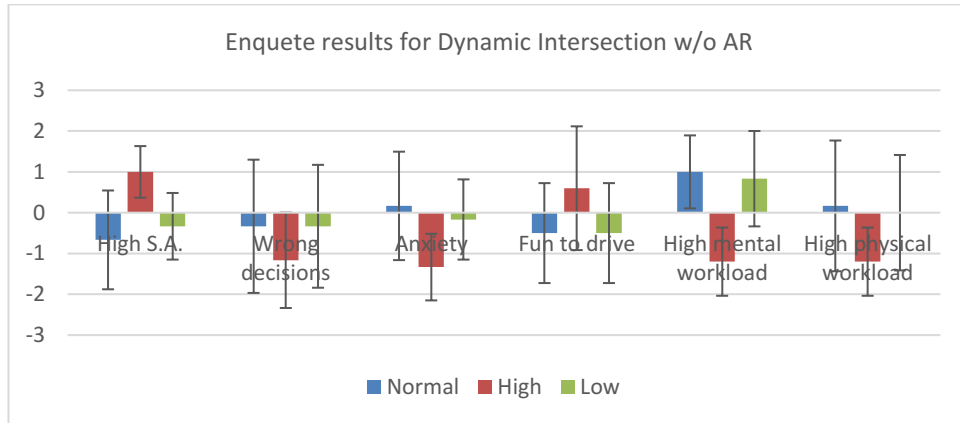


Figure 5.13: Enquete results for dynamic intersection without AR.

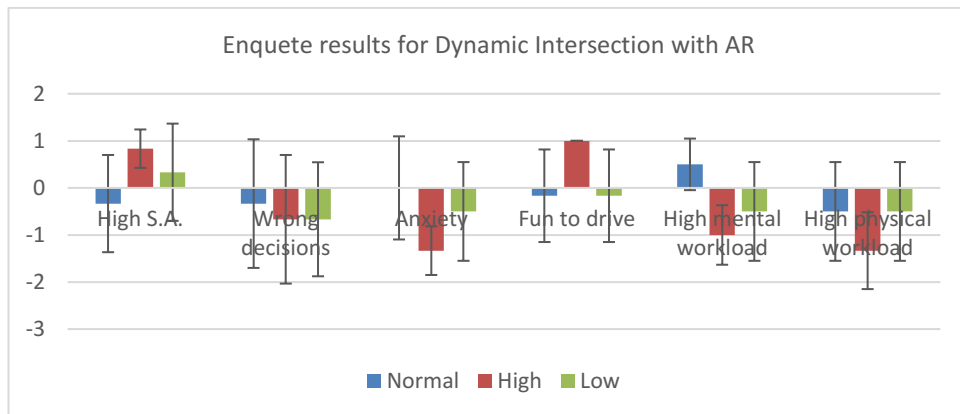


Figure 5.14: Enquete results for dynamic intersection with AR.

### 5.3.3.1 Static intersection

Four out of six participants made all decision for the road to take at the static intersection correctly. There are too few cases of wrong decisions to draw conclusions. The trend for the range of head rotation is similar to dynamic intersection and shows a decrease for the high viewpoint.

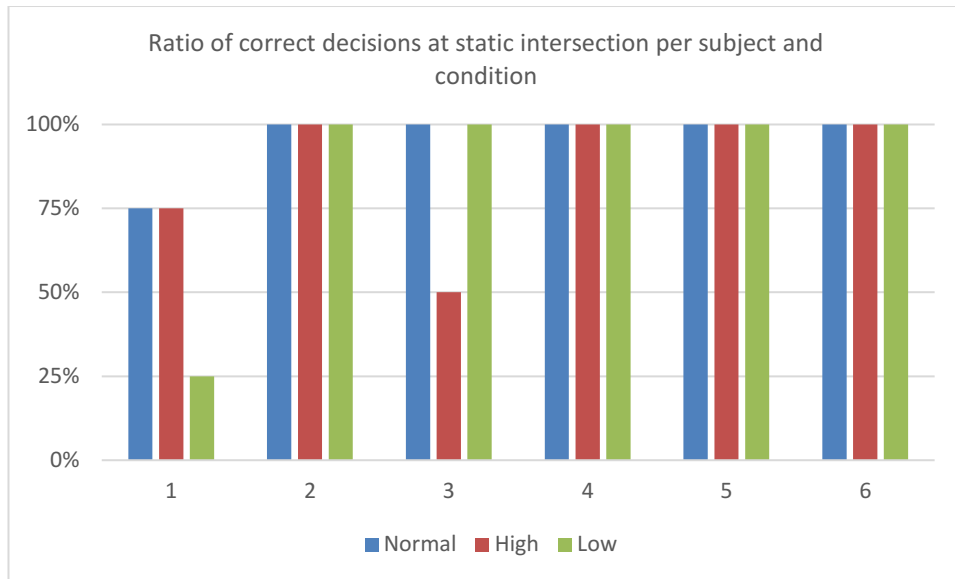


Figure 5.15: Ratio of correct decisions.

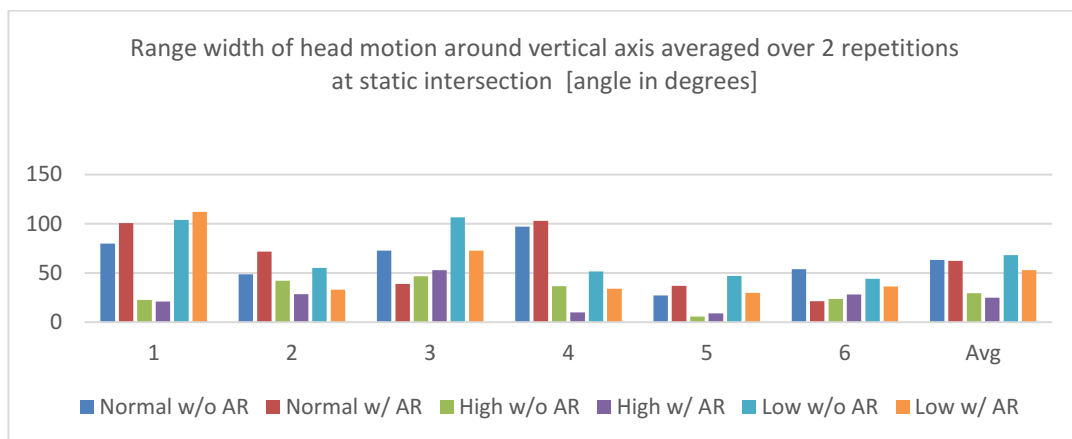


Figure 5.16: Range of head motion around vertical axis at static intersection.

The enquete results for static intersection differ from all other sections and show an advantage of normal and low viewpoint over high viewpoint in all aspects.

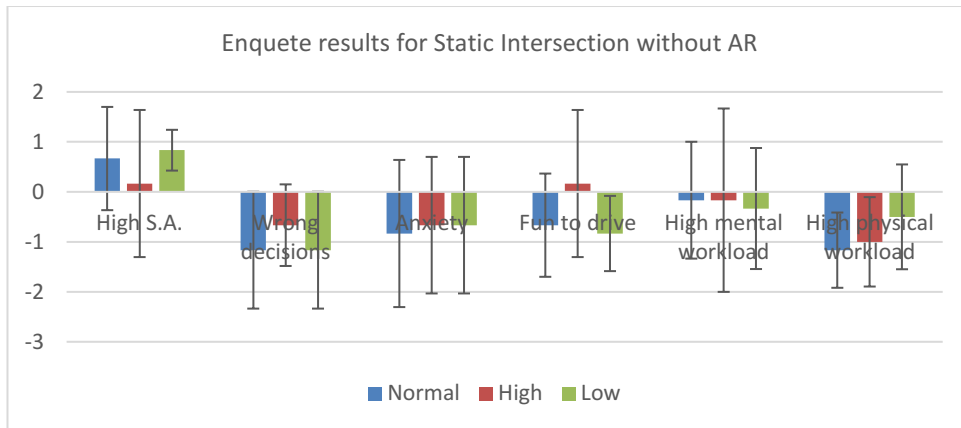


Figure 5.17: Enquete scores for static intersection without AR.

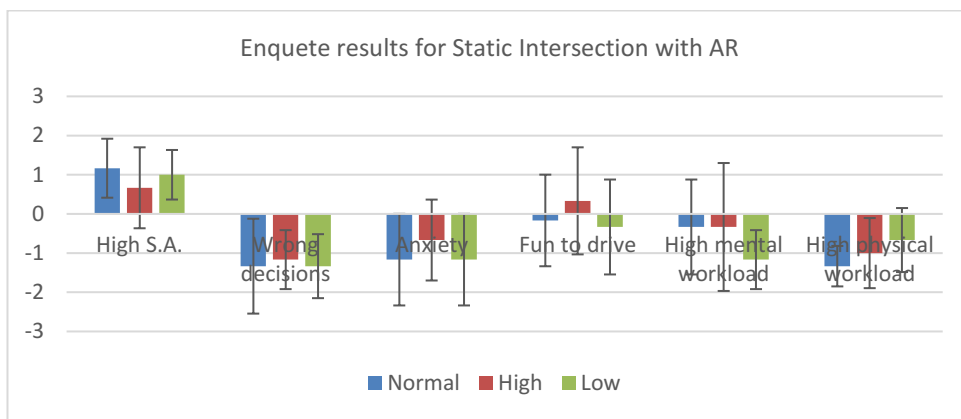


Figure 5.18: Enquete results for static intersection with AR.

### 5.3.3.2 Curve and Crank

The average distance driven off-road showed different trends for curve and crank sections. In the curve, high viewpoint was similar to normal viewpoint on average but showed larger variation, while low viewpoint was more difficult for participants. For the crank, high viewpoint was better than normal viewpoint although it also showed larger variation, while low viewpoint was midway between normal and high viewpoint. The performance per participant mirrors these trends. High viewpoint was most likely to complete without veering off road.

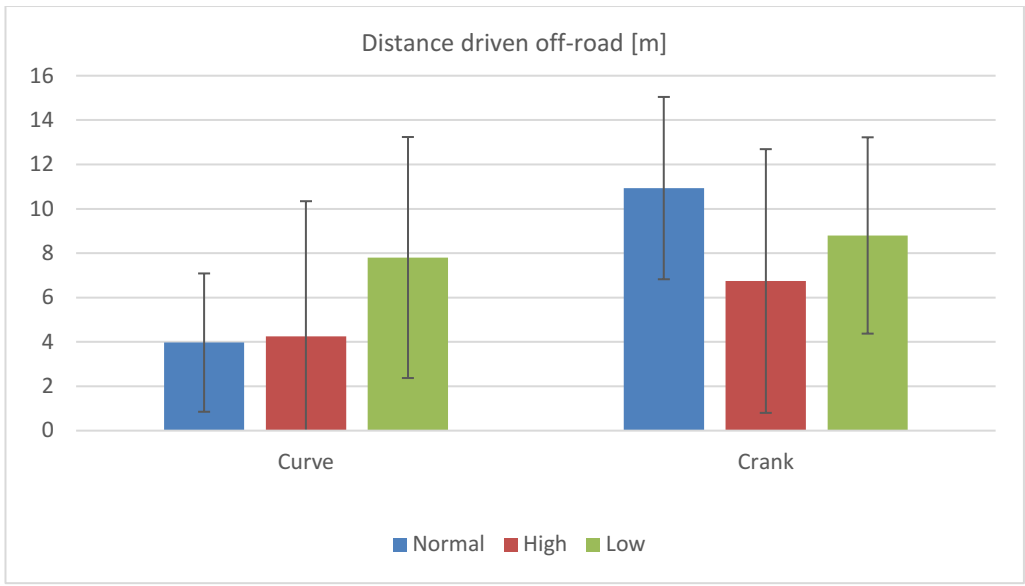


Figure 5.19: Average distance driven off-road.

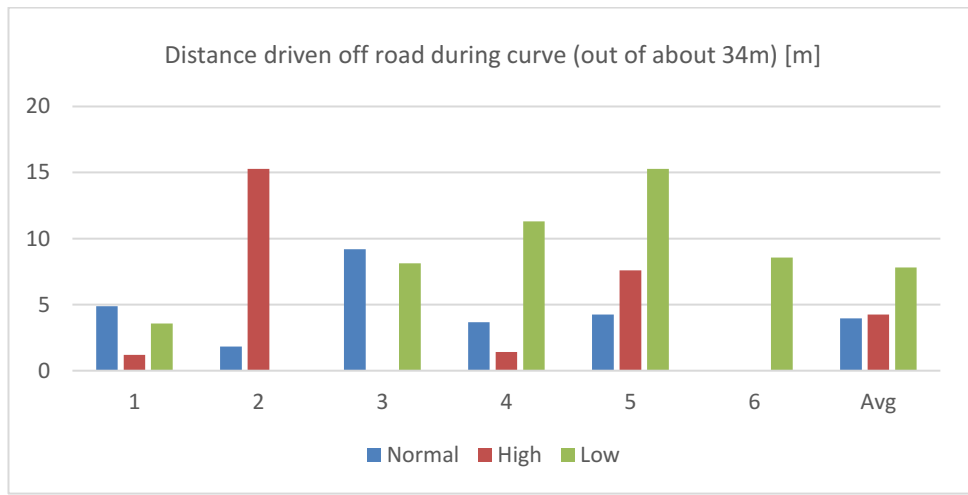


Figure 5.20: Distance driven off road during curve per participant.

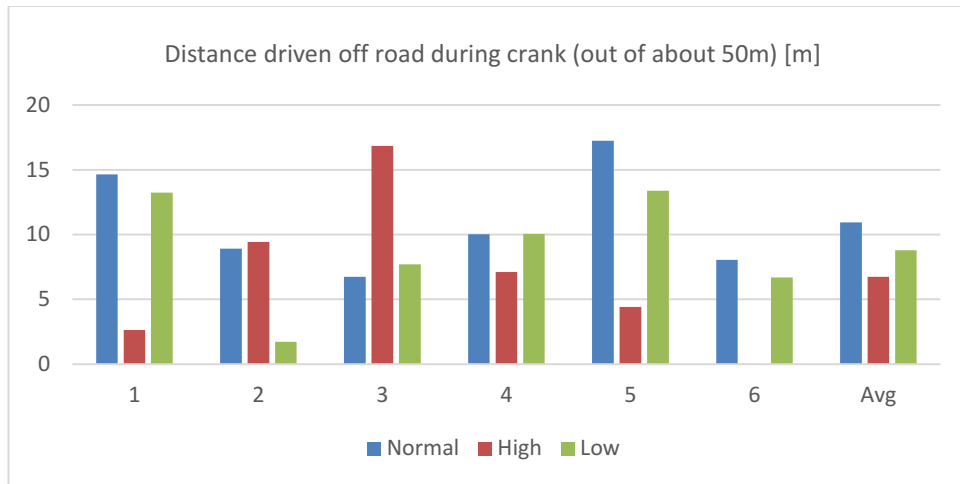


Figure 5.21: Distance driven off road during crank per participant.

Despite these differences between curve and crank, enquete results are similar for both and show an advantage for high viewpoint in all aspects.

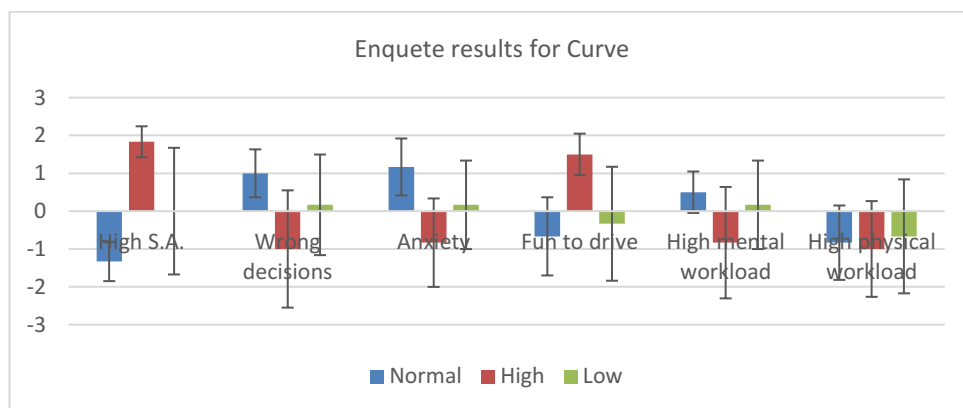


Figure 5.22: Enquete results for curve

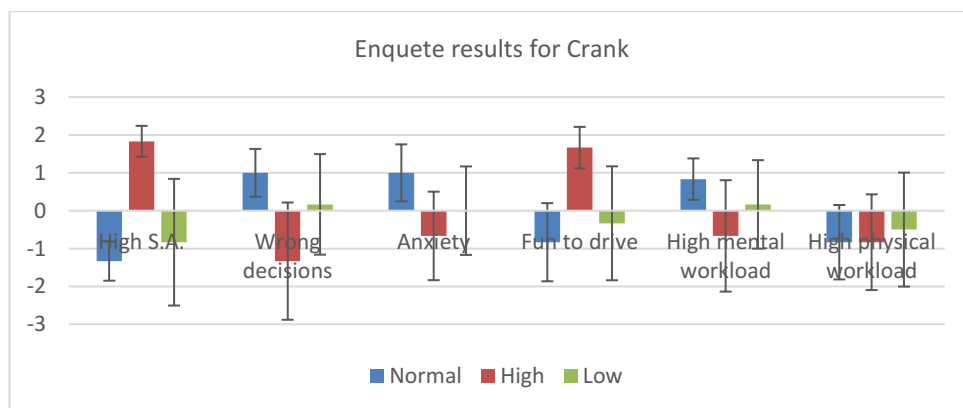


Figure 5.23: Enquete results for crank

### 5.3.3.3 Parallel parking and parking exit

Duration spent for parallel parking show a trend similar to the duration for dynamic intersections.

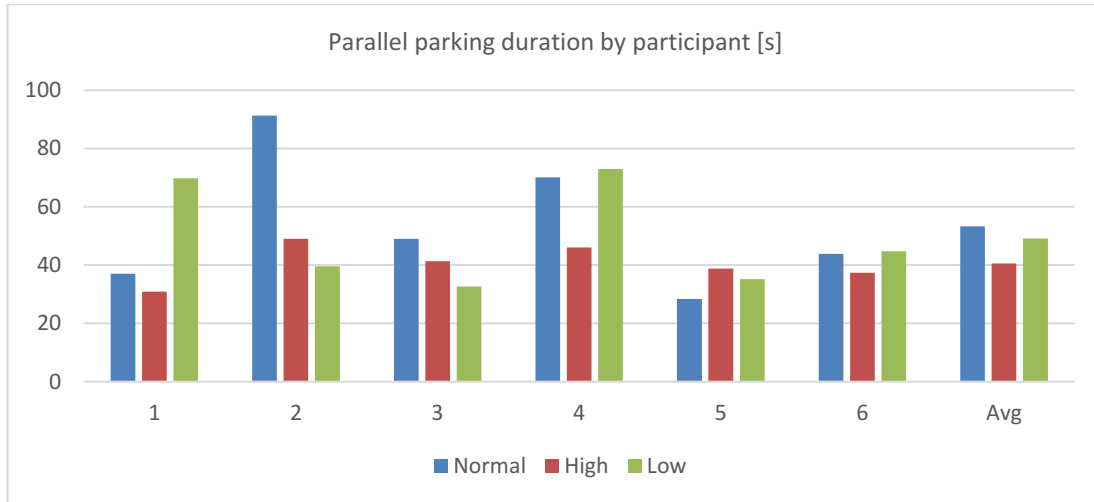


Figure 5.24: Duration of parallel parking.

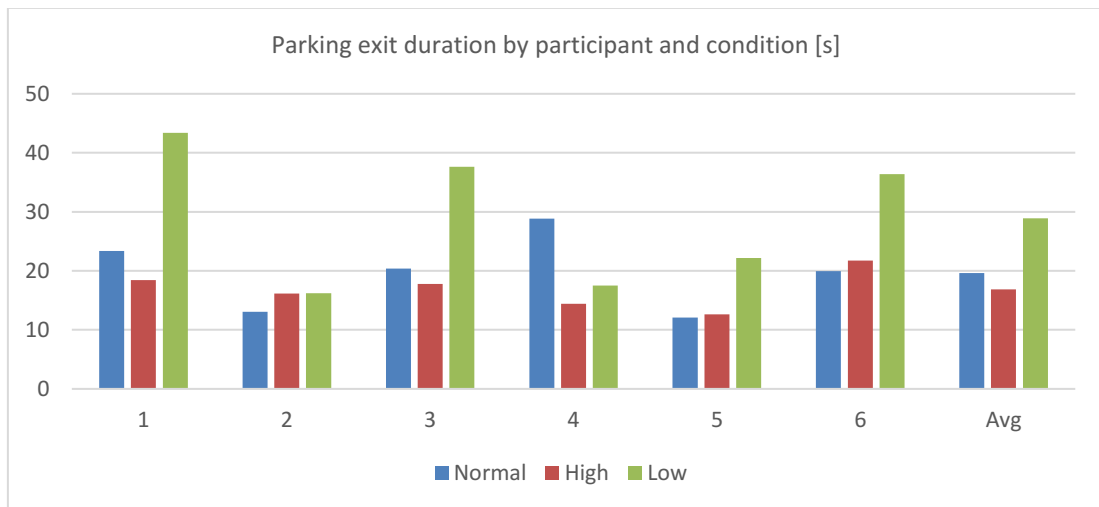


Figure 5.25: Duration of parking exit.

The number of direction changes, i.e. switching either from forward to backward driving or from back to forward, was, for parallel parking, on average smaller for high viewpoint compared with normal viewpoint with the low viewpoint in between. For parking exit, normal and low viewpoint were about equal with the high viewpoint showing more direction changes on



average. These trends persist when looked at a per participant basis. (Figure C.17, Figure C.18)

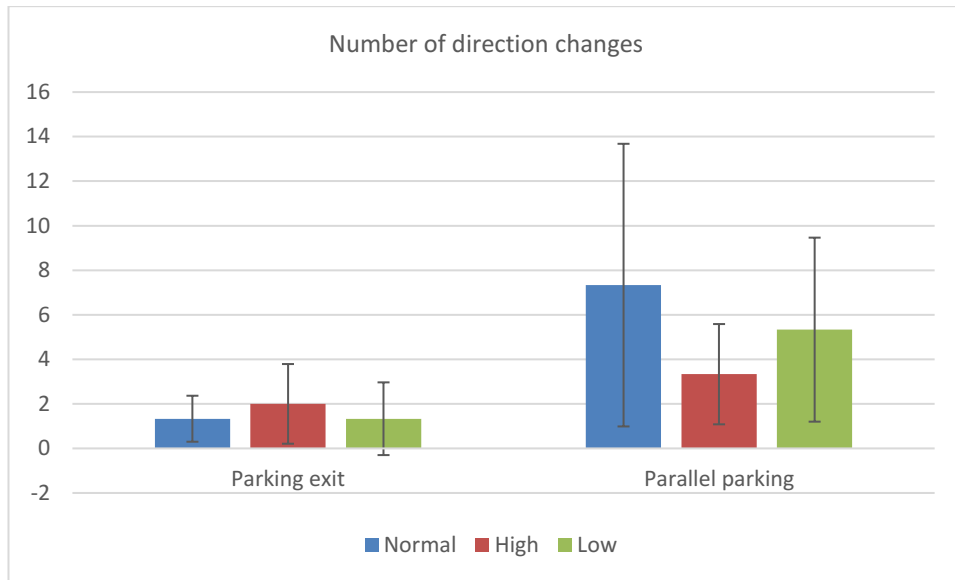


Figure 5.26: Number of direction changes.

Enquete results for parking exit and parallel parking both show similar advantage for high viewpoint in all aspects.

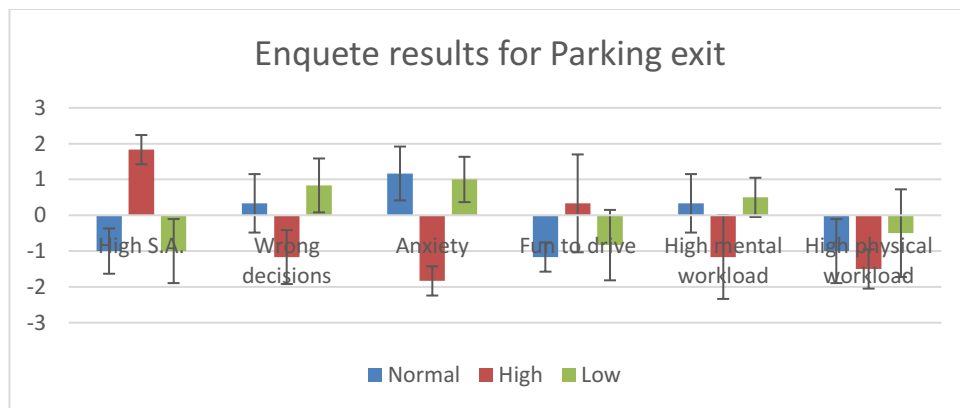


Figure 5.27: Enquete results for parking exit.

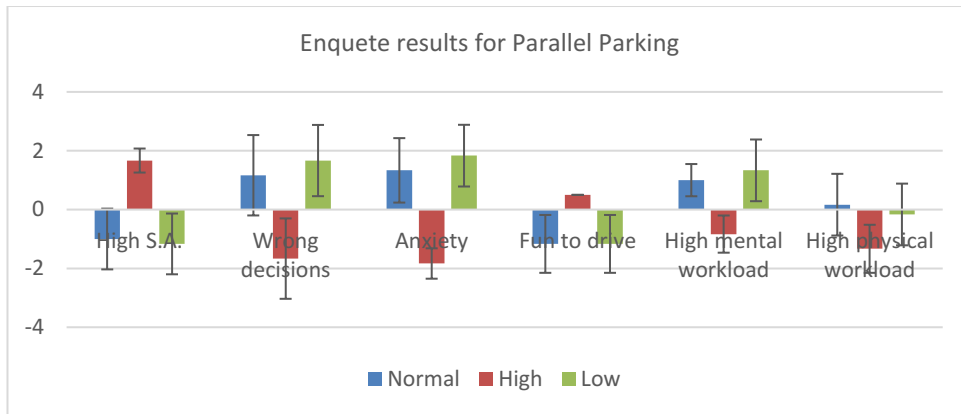


Figure 5.28: Enquete results for parallel parking.

### 5.3.3.4 AR

When comparing AR against no-AR at both intersections, we see a slight advantage in average duration for AR. This trend also holds for the durations on a per participant basis. (Figure C.19, Figure C.20)

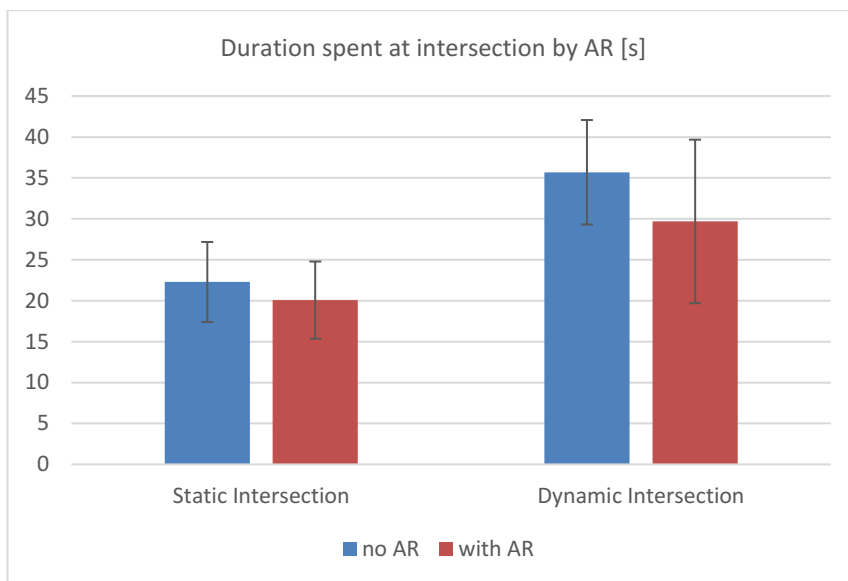


Figure 5.29: Duration of intersection by AR.

A comparison of enquete scores of AR against no-AR, on average, there seems to be a slight advantage of AR with only situation awareness and mental workload showing a difference larger than 0.5 on the rating score. All scores show large variability between participants and the only score with small variability is that mental workload is low with AR. A more detailed comparison on a per participant basis (but averaged for static and dynamic intersection) shows more specific results (Figure C.21 to Figure C.26):

- Participant 1 indicates no difference between AR and no-AR except for giving AR slightly worse scores by about 0.5 than no-AR in the high viewpoint condition.
- Participant 2 shows no difference between AR and no-AR.
- Participant 3 generally scores AR better than no-AR by 1 to 2 scores in all aspects and for all viewpoints.
- Participant 4 scores AR slightly better than no-AR in all aspects but not for the high viewpoint.
- Participant 5 scores AR slightly different from no-AR by around 0.5 but no clear trends.
- Participant 6 sees effect of AR for high viewpoint only as contributing to driving enjoyment. For normal viewpoint there are slight differences between aspects but no clear trend. For low viewpoint, AR contributed positively to S.A. , anxiety, fun to drive and mental effort by more than 1.

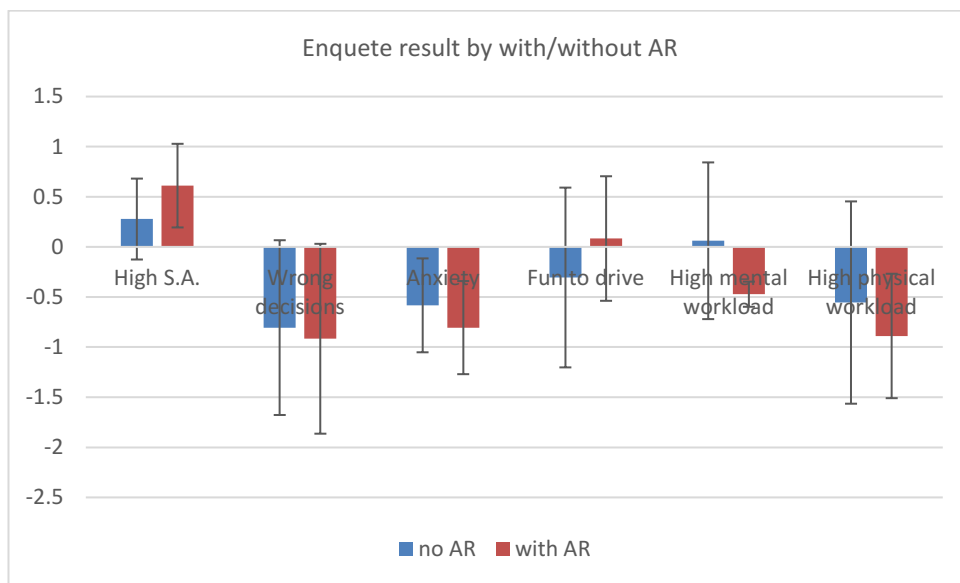


Figure 5.30: Enquete results comparing AR against no-AR.

#### 5.3.4 Other observations

Motion sickness was an issue for some subjects. If an issue, it affected later conditions more.

Many subjects did not move their body to change the blind area caused by their car's A pillars and door mirrors when waiting for an opportunity to cross the dynamic intersection. Questioning them after the experiment

---

indicate that this might have been a limitation of experimenting using HMDs which caused participants not noticing that moving their body would change their blind spot.

Understanding and utilizing AR in accordance with their intended meaning and usage seemed to have been difficult for some. While the virtual barricade in the static intersection was simple and the fact that it was missing in one of the two other directions did not cause any questions, the virtual safe area indicator in the dynamic intersection condition was sometimes confused as an additional danger with which one should not get in physical contact, or as indicating safety for cross traffic in both directions.

The cross-traffic pattern was implemented as a cyclic repetition of predefined, constant distances between cross traffic. While the pattern itself was most likely not noticed, some subjects noticed that cars would appear when the pattern reset.

## 5.4 Prototype Implementation

To answer to the second research question, i.e. whether manipulation of viewpoints is realizable using current technologies, we implemented a drivable prototype.

In order to enable left turns of up to 90 degrees using the actual driver eye-point as viewpoint in addition to driving straight forward, a side display was added on the left side of the frontal display to cover a geometric field of view of about 70 degrees to the left. The geometric field of view to the right was limited by the cockpit layout that did not permit adding a display on the right side. Camera arrays were placed forward of the displays to feed the multiple viewpoints of the auto-stereoscopic displays. Using auto-stereoscopic displays, drivers did not have to wear stereo glasses.

Manipulation of the viewpoint was implemented by real-time viewpoint transformation of the surroundings fused with information from detailed 3D maps. This 3D reconstruction was implemented using a Velodyne rotating laser range finder to obtain a 3D point cloud of the surroundings at a refresh rate of 10 Hz, which was then converted into 3D meshes. The image streams of multiple cameras including a Point Grey Ladybug spherical camera were texture mapped onto the meshes using projective texturing technique implemented in a custom OpenGL shader that utilized the Open Scene Graph library. The system automatically chose an appropriate point of view depending on the current driving situation triggered by locations on the map.

Semantic enhancements were implemented as CG augmentations to the image and included driving rules like speed limits, road geometry like the drivable area between the road boundaries or the left curb and the center line, the permitted paths to take at intersections, predicted vehicle trajectory at the current steering angle, and seeing the road on the other side through occluders. through buildings.

Sensory enhancement was implemented using highly sensitive night vision cameras.

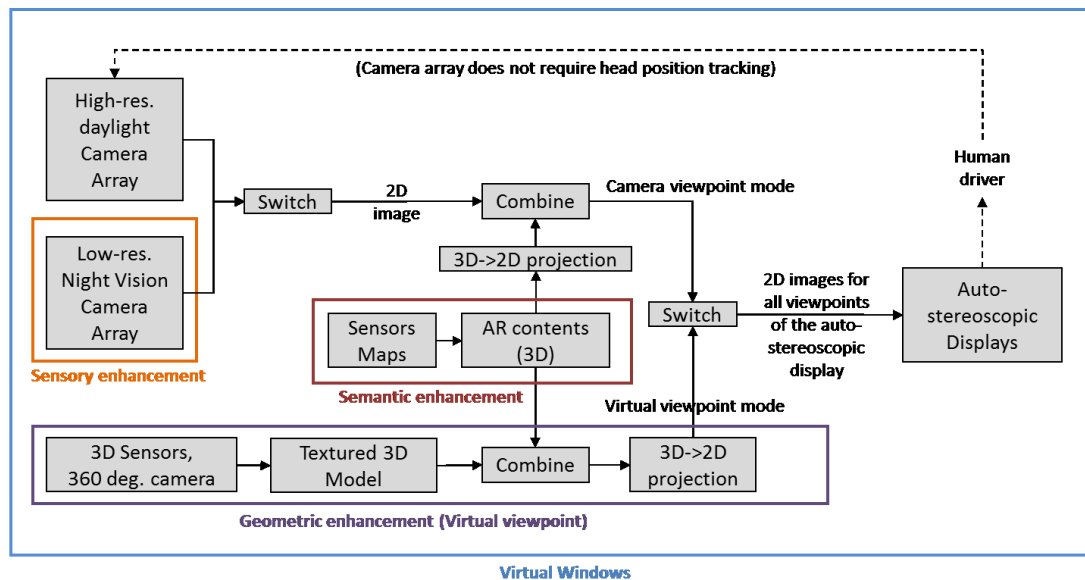
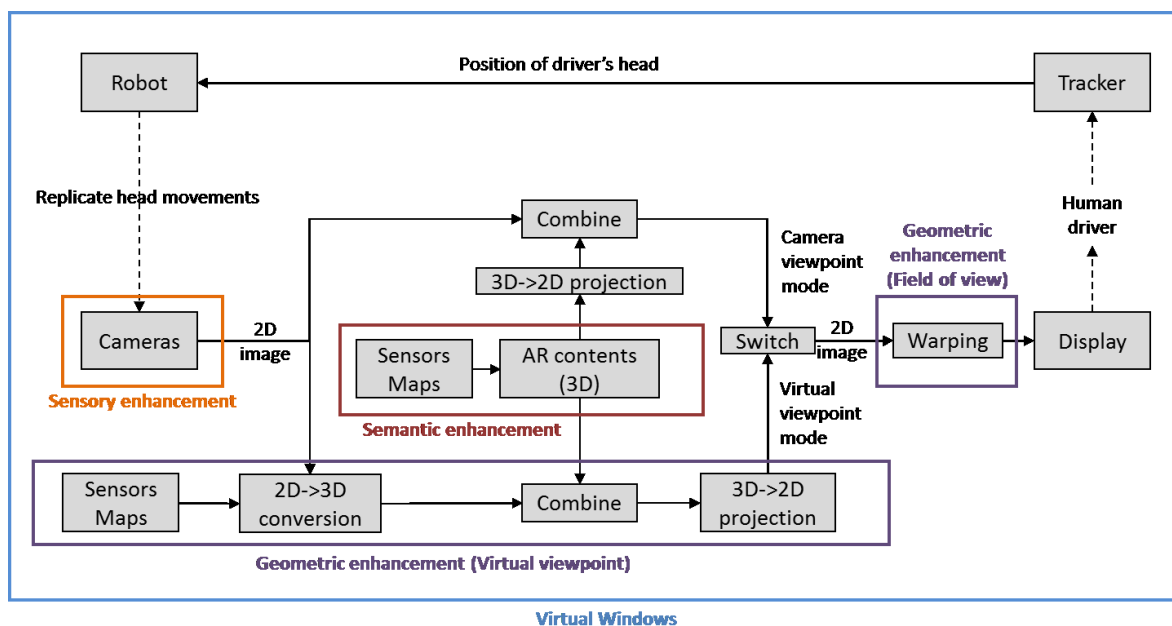


Figure 5.31: System architecture of the prototype.



---

Figure 5.32: Generalized architecture adding capability to manipulate distortion in addition to point of view.



Figure 5.33: Exterior of the experiment vehicle (Virtual windows)



Figure 5.34: Interior of the cockpit with the two auto-stereoscopic displays. (Virtual windows)



Figure 5.35: Life-size virtual rearview with a visualization of the expected trajectory. The image is left-right reversed in order to match the visual

---

image to the accustomed steering rotation direction. The image in the left upper corner shows the position of the experiment vehicle (silver minivan) relative to the surroundings. (Geometric enhancement)

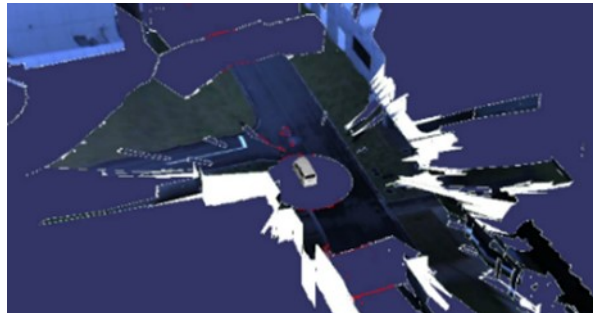


Figure 5.36: Screen capture of the front display showing a bird view image from a high viewpoint above the car. The camera images from the surround camera have been combined with 3D geometry of the surroundings captured by the Velodyne lidar. (Geometric enhancement)

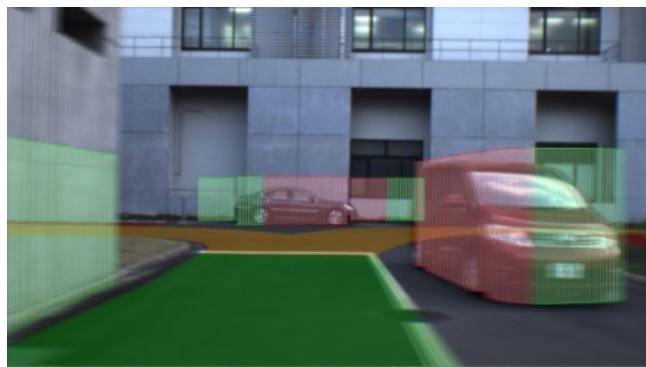
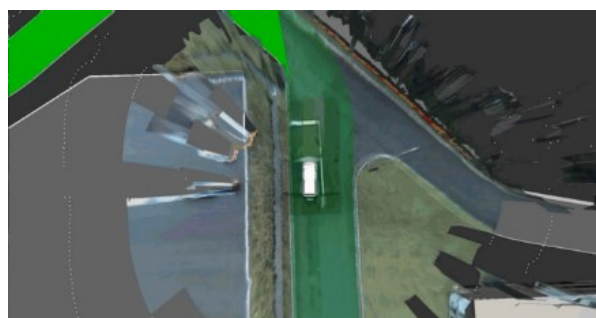


Figure 5.37: The area and directions permitted to drive are colored in green (indicating area of own priority over other traffic) and orange (indicating that other traffic has priority). Three-dimensional obstacles within the road boundaries are marked in red. (Semantic enhancement)



---

Figure 5.38: Virtual topview of an intersection with test drive course marked in green. (Combined Geometric and Semantic enhancement)

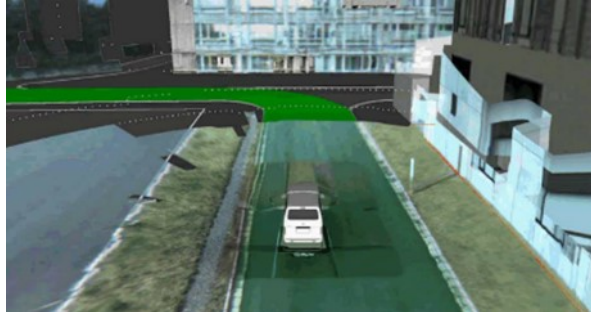


Figure 5.39: Virtual perspective birdview that combines camera image and 3D map and a CG representation of the own car. (Combined Geometric and Semantic enhancement)



Figure 5.40: Close perspective birdview with a transparent representation of the own car and only the wheels and steering visible to help accurate maneuvering relative to the curb on the inside of the left turn. (Combined Geometric and Semantic enhancement)

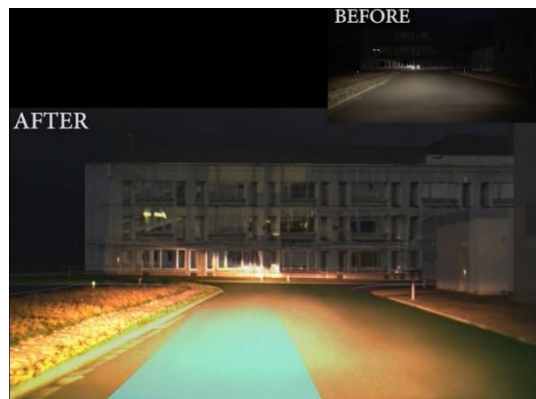


Figure 5.41: Image from a high sensitivity camera combined with augmentation of buildings, path and trajectory.



#### 5.4.1 User feedback from test drives

This system was driven on roads within our research facility by selected Nissan executives and employees and verbal feedback collected. While the overall opinion was somewhere between interesting, intriguing and useful, we also identified some possible negative side effects: selection of a single optimal viewpoint might not always be possible, transitions and change of viewpoints can be confusing, third person perspectives might reduce immersiveness and feeling of presence and danger, the changed optical flow could make intuitive reaction and skill based action more difficult.

Table 5.5 : User feedback about viewpoint

Item	Feedback
Driving situation dependent views	Useful and better situation awareness. In some cases, selection of viewpoint was different from expectation and confusing. Might requires time to get skilled with new views.
Backward view	Convenient and subjectively better SA. Some subjects reported motion sickness, issues with abrupt switching and unexpected vehicle path due to effectively steered rear wheels.

Table 5.6 : User feedback about AR

Item	Feedback
AR tracking precision and delay	Enough precision for slow, straight driving, but delay perceivable when turning at intersections.
AR predicted vehicle path at current steer angle	Useful to understand the direction the car is moving, especially in third person viewpoint. A long path at constant radius can be irritating in curves as it goes off-road.
AR Visualization of road and road boundaries hidden behind occluders	Useful for planning ahead, but semi-transparent objects can lead to misjudgment of presence and distance of those objects.
AR Virtual traffic signs	Useful as a reminder. Could be improved when combined with driver monitoring or reacting to overspeed.
AR Route navigation	Useful and makes following a route dead simple. Coloring unclear at road boundary and objects. Might interfere with driving skills.

---

## 6. Replicating Motion Parallax from Head Motion<sup>3</sup>

### 6.1 Objective

Human vision obtains distance cues from many sources. While it is not impossible to drive with one eye closed or without head motion, driving with both eyes using binocular stereo cues is usually more comfortable, and using body motion is often useful to understand the spatial configuration at close distance. While motion cues work particularly well for close objects, the absence of change in perspective of far objects can be a cue for distance of far objects.

Motion parallax from head motion are small changes in viewpoint depending on head motion. Motion parallax from head motion is present when directly seeing the surroundings and known to be a relatively strong depth cue but not replicated in many indirect vision systems. What elements of perception could replicating motion parallax from head motion improve that are otherwise degraded in an indirect vision system, and how should such a system be implemented?

### 6.2 Design and Implementation

There are mainly two approaches to produce seamless motion parallax. The method we chose is a multiple DOF robotic stereo camera designed for use in Telexistence applications that tracks the head motion of the driver. Alternatively, a multiple camera array could be used, and images interpolated for arbitrary viewpoints between the camera locations in the array. The former method has the merit of higher resolution and less artifacts by not relying on interpolation. In order to exactly match interocular distance with the subject, the distance between the stereo camera pair should be adjusted for each subject. For simplicity, our current system uses a fixed distance. Concerning latency, while the former method will show some motion latency as well as robot tracking inaccuracies, the latter method will show processing latency and inaccuracies from interpolation.

---

<sup>3</sup>T. Yanagi, C. L. Fernandez, M. Y. Saraiji, K. Minamizawa, S. Tachi and N. Kishi, "Transparent Cockpit Using Telexistence," *IEEE Virtual Reality Conference*, pp. 311-312, 23-27 March, 2015, Arles, France.

---

The robotic head system consists of a XY robot and a custom made 3 DOF robotic head. We decided not to implement Z motion because the driver's head does not move much in the vertical direction and a preliminary driving simulator experiment showed that ignoring Z motion did not have a significant effect. The head uses two web cameras fitted with 128.5 degrees wide-angle lenses as a stereo camera pair placed 65mm apart so that it matches the average human Inter-Pupillary distance in order to create correct distance and depth perception. The completed robot, placed in front of the back-side of the display and above the engine hood. In order to generate control commands for the robot, the driver's head motion is captured by a motion tracking system.

An initial calibration matches the default head position of the driver to the default position of the robot. After the calibration, the x, y position as well as pan, tilt and roll of the robotic head moves so that the stereo camera pair maintains constant relative position and direction to the drivers' head. The XY robot and the head communicate at 200Hz cycle speed after filtering rugged motion with a digital low pass filter. The driver's head motion is captured from a motion tracking system (Model: Opti Track Duo) and converted to the motion of a 5 DOF robotic head. The robotic head system consists of a XY robot (Model: IAI LSA-S6SS, LSA-S8HS series) and a custom made 3 DOF robotic head with pan, tilt and roll motion.

A Nissan NV200 minivan was used as the experimental vehicle to implement a camera-based vision-by-wire system without camera arrays and buffering. The front windshield was replaced by a large, stereo-capable 60-inch LCD display. The stereo image pair is mapped onto a virtual projection screen and the resulting image is shown to the driver wearing an active shutter 3D glass (Model: Sharp AN-3DG20-B) to which retro-reflective markers for the motion tracking system were added. Therefore, when the driver moves in x, y direction as well as pan, tilt, roll the robotic head in front of him moves accordingly to give the exact same point of vision.

The stereo cameras are placed 1560mm in front of the driver's head. An initial calibration is used to calculate the initial drivers position and accordingly the robotic head is moved to maintain the distance. The left and right eye cameras are placed 65mm apart so that it matches the human Inter-Pupillary distance. (IPD). For tracking the users head motion, 3 trackable markers (11mm diameter, Retro-reflective markers) were placed on the sides of the 3D glass. Then a rigid body is composed with the 3 visible markers at a time. Once the center of gravity point is determined, the pivot point was shifted back towards the driver's head at 80mm in order to pivot around the head center.

The FoV of the initial camera module was 74 Degrees and the vision sensor is a 16:9 format that provides an aspect ratio of 1.77. This leads to a working environment of 35 Deg (Vertical) and 65 Deg (Horizontal) FoV. However, as shown in Figure 2.4, the ideal FoV needed to provide a true active wind-screen experience it has to be 120 Deg on Horizontal and about 50 Deg Vertical Field of View. In order to satisfy the above conditions, a special wide-angle lens was mounted to the camera modules where the wide conversion ratio was 0.5. With this, the captured vision's FoV was increased to 128.5 Deg (H), 72.5 (V) Deg. This will effectively provide a close but not ideal FoV for the requirements.

The XY robot and the Head communicate over two dedicated hardware RS232 buses at 115200bps. The robot commands are processed at 200Hz cycle speed where as rugged motion is filtered with a digital low pass filter.

While the initial system was designed with two linear motors allowing for lateral and longitudinal motion and located above the eye point height with a 3-axis robotic camera hanging downwards implementing 5 axes of motion, we observed that vibrations from the engine of the car resulted in vibration of the image caused by lack of rigidity relative to the length of the arms and minuscule play in the rotational axes of the robotic camera. In order to reduce the magnitude of the vibrations, we changed to a design with a single lateral linear motor located below the eye point and a robotic camera without rotation around the roll axis and 3 axes of freedom in total.

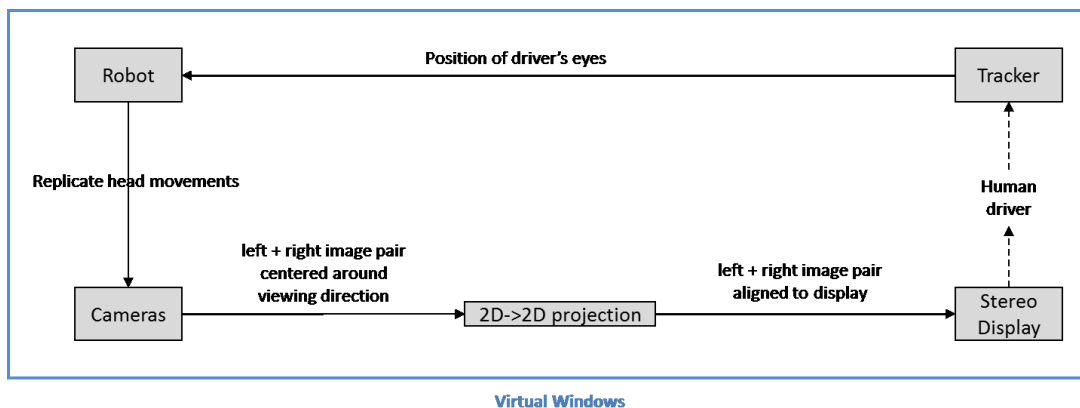


Figure 6.1: System architecture.

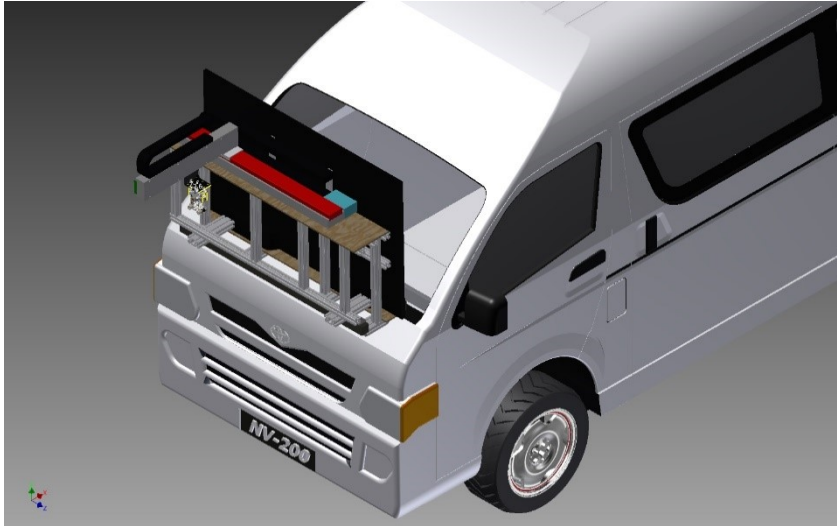


Figure 6.2: Initial design for the Virtual Window implementation using a robotic camera in front of a large display that replaces the windshield.

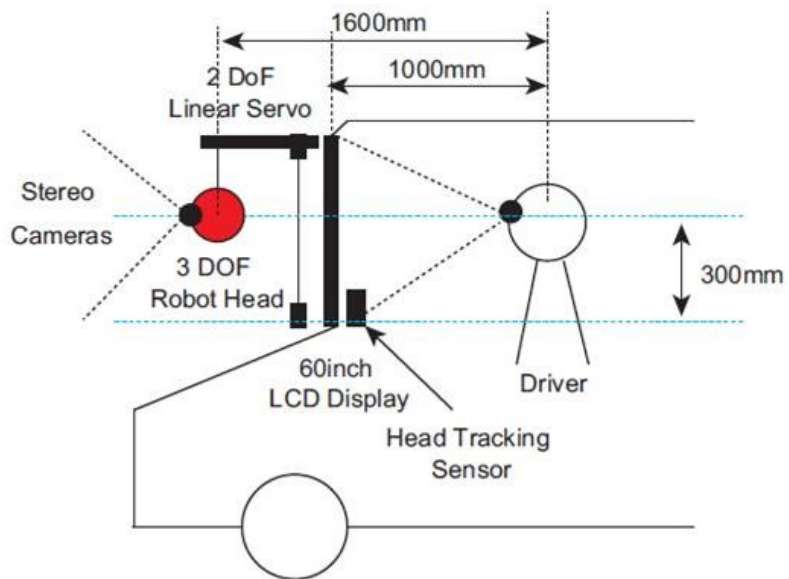


Figure 6.3: Side view of the system layout

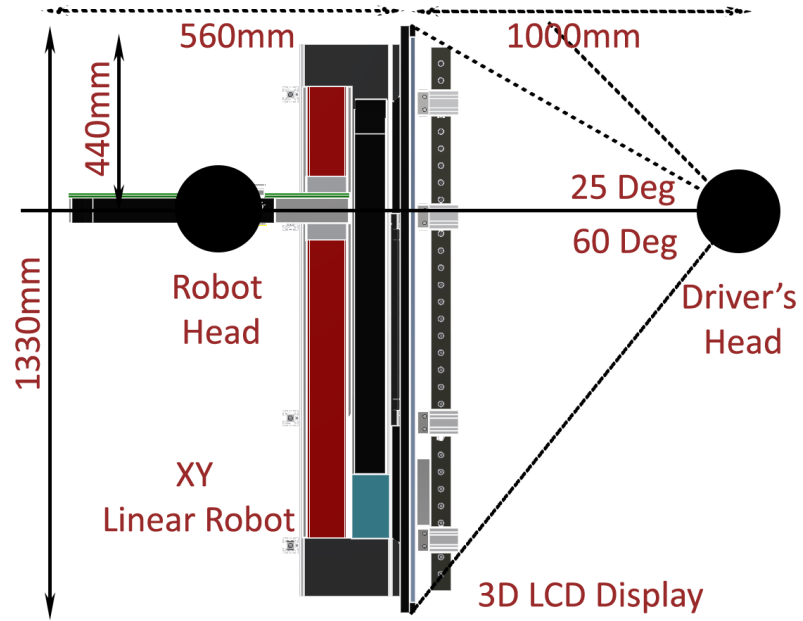


Figure 6.4: Top view of the system layout

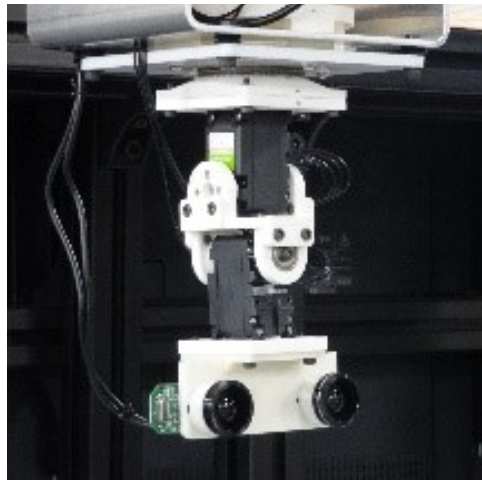


Figure 6.5: Image of the initial version of the robotic camera



Figure 6.6: Image of the cockpit interior with view to the display



Figure 6.7: Exterior of the initial version of the experiment vehicle with the robotic camera hanging downwards from the linear motor mounted on top.

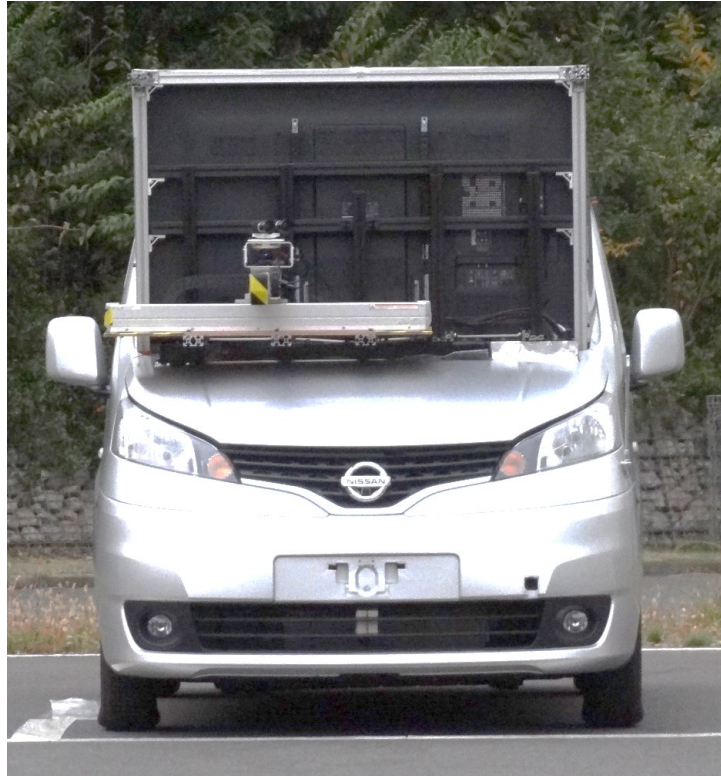


Figure 6.8: Front view of the final version of the experiment vehicle with an upward robotic camera above the linear motor.

## 6.3 Evaluation

### 6.3.1 Objective and Hypothesis

The aim of this experiment was to test the hypothesis that a combined stereo image and motion disparity condition would result in a more precise and accurate perception of both space and speed than a stereo image without motion condition and a monocular image condition, which would ideally show in a more precise driving performance through a test course.

### 6.3.2 Participants and Experiment Procedure

N=4 student participants first drove an unmodified standard NV200, whose field of view had been masked by black tape to match the field of view of the Virtual Window implementation, through a test course to establish baseline performance before repeating the task in the three experiment conditions, which were a monocular image, binocular stereo and combined binocular and motion stereo. Given the small number of subjects, we decided not to randomize the order. The vision-by-wire system replaced only the front window, and the side windows and mirrors were masked in order to have the driver rely only on the view obtained through the vision-by-wire system. This unluckily made it difficult to maneuver through right turns as there was simply not enough field of view to the right. Because of the limited field-



---

of-view of the prototype, 45-degree turns to the left were the maximum turning maneuvers that we deemed safe. The course consisted of two 45-degree turns to the left, an obstacle avoidance maneuver simulating avoiding a parked vehicle, and stopping in front of an obstacle to measure the effect of differences in spatial perception. We placed obstacles on both sides of the course and ahead of the final stop line in order to provide three dimensional visual stimuli about the course, its turns and the final stop location. When judging the point where to start turning left after having passed the left turn corners, subjects had to rely on their intuition of having seen the corner in the frontal vision-by-wire system and the approximate distance that the car had progressed after that moment.

Subjects were asked to drive safely along the course without leaving the area between the road boundary marked by the white lines and to stop exactly at the final stop line.



Figure 6.9: View of the test course from a nearby building.



Figure 6.10: View of the test course from eye height.



Figure 6.11: View of the test course from the normal car used in the base condition. The bottom of the windshield and the upper left corner were

---

masked with black tape to limit the field of view to the same area as provided by the Vision-by-Wire implementation to ensure that the experiment results are not influenced by field-of-view and comparable.

### 6.3.1 Results

All participants were able to complete all experiment conditions without accidentally touching the obstacles and with only minimally veering outside the white boundary of the course.



Figure 6.12: Images of an experiment run.

---

When looking at the trajectories of the participants, they all showed the same tendency. The base line condition appeared naturally centered on the test course, while the test conditions led to early left turns, with monocular condition worst and almost touching the road boundary, and both stereo and combined stereo and motion conditions about equal and between base line and monocular condition. The figure below shows the trajectory of one subject.

The offset of the trajectories of the stereo and combined condition is about equal to the horizontal distance between the eye position of the subject and the cameras and showed that participants were not able to or did not intend to correct their driving for the change in eye point. While the absence of any cues about the camera position relative to the own car was one reason, this shows that this offset is of concern for exact maneuvering. It is unclear how much it will affect driving at higher speeds where the turning radii are less tight, but we assume that it will be less.

Monocular condition was worse and indicated that the lack of binocular cues degraded spatial perception of depth.

The absence of a clear difference between the binocular condition and the combined binocular and motion stereo condition may indicate that ...

The results are discussed in chapter 8.



Figure 6.13: Image of a situation where a left turn was initiated a bit too early.

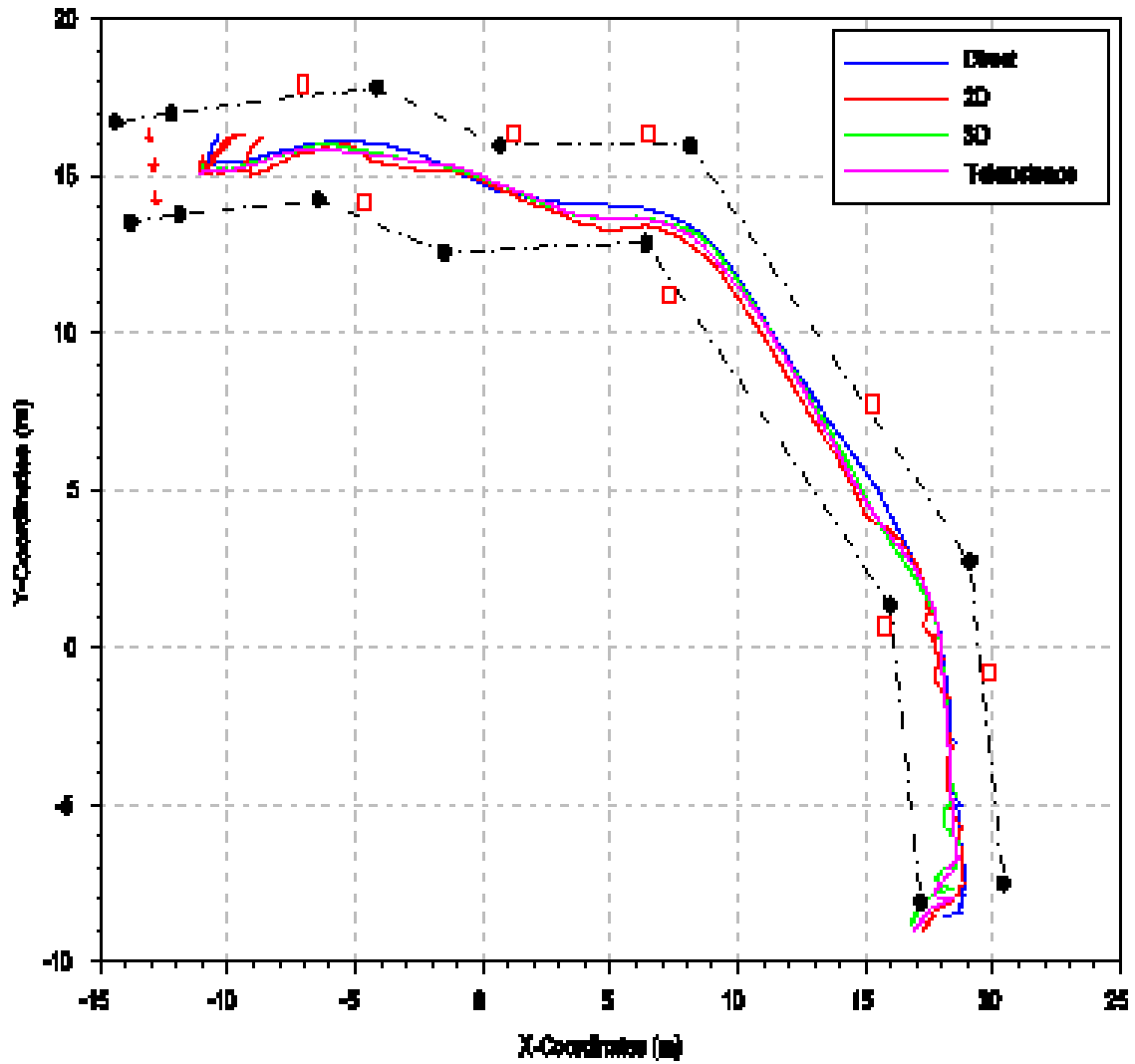


Figure 6.14: The trajectories of the four experiment conditions of one of the participants. The colored lines indicate the trajectory of the center of the rear wheel. (Black dotted line: course boundary marked by white lines, red boxes: obstacles, red dotted line: stop line at the end of the course.)

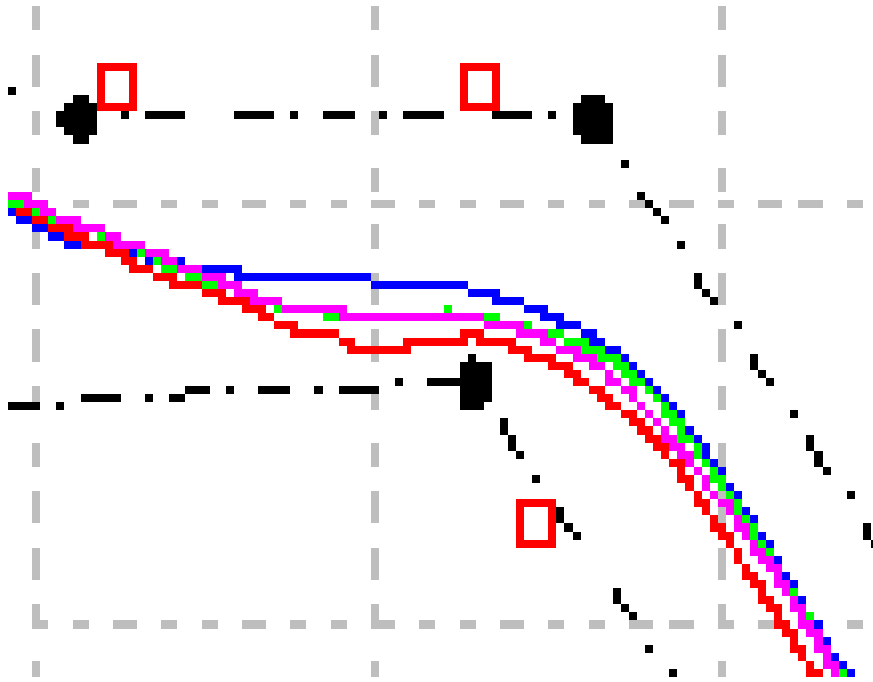


Figure 6.15: Enlarged top view of the trajectories at one of the left turns, approaching from the bottom right and leaving towards upper left. The colored lines indicate the center of the rear wheels. (Blue line: base condition, red: monocular, green: binocular, pink: combined binocular and motion stereo.)

---

## 7. Discussion

This thesis studied automotive indirect vision cockpits as an example application for investigating the possibility of indirect vision systems to improve human spatial awareness. Electronic indirect vision systems provide a unique opportunity to manipulate the projection method used to map the three-dimensional surroundings onto one or more two-dimensional displays. Other changes, like improving the signal-to-noise ratio or adding supplementary information by augmented reality, are possible with indirect vision systems but also with other approaches like semi-transparent head-up displays.

Projections map three-dimensional space into two dimensions for presenting space on two-dimensional displays. Projections are therefore a core part of any vision system that handles three dimensions and justifies a closer investigation. A digital indirect vision system can implement projections that are optimized for human vision without constraints from the optical design of lenses or mirrors. For example, projections that are not axially symmetric can be used.

Projections largely fall into two categories: parallel and perspective. The former is used for example in technical drawings and require large display space to cover the field of view needed for general driving. Perspective projections have one or multiple vanishing points and can cover the entire space to the infinitely distant vanishing points on a small display and are therefore suited for displaying the surroundings while driving.

Perspective projections can be described by its distortion and the point of view which is defined by the location and direction of view.

While a rectilinear projection from the actual eye point produces a image closest to the visual experience of seeing the surroundings from that view point, that is not necessarily the most effective projection resulting in the best spatial awareness, the best task performance, or the best experience. To make an even stronger case, the way humans perceive the real world surrounding them using their own eyes is not necessarily objective and true and not necessarily most effective.

This thesis is based on the hypothesis that, similar to map projections that transform locations from the surface of a sphere or an ellipsoid to locations on a plane, projections of three-dimensional space into two dimensions for use in indirect visions systems should be designed and chosen depending on the task and its requirements. Indirect vision system can therefore be

---

thought of as a tool to optimize or correct the spatial awareness of the surroundings that humans are natively capable to obtain using their own eyes by direct sight.

Given the task specificity of optimizations, any concrete investigation needs to focus on a specific task, even if the results should hopefully be generalizable. This thesis focused on car driving as a popular task that requires a high level of spatial awareness and allows the system to be implemented as part of the car cockpit without the user having to carry the system including sensors, image processing system and displays.

The effect of distortion on spatial awareness was studied using rearward vision for lane changes as a test case. Combining the visual requirements with known properties of human visual perception led to the development of a novel projection method that combines characteristics from perspective and parallel projections and was used to implement a rear-view system that significantly outperforms conventional mirrors and camera systems. This can also be seen as an attempt to challenge the notion that a distortion-free representation is ideal.

The effect of point of view on spatial awareness was studied in a simulated sequence of typical driving tasks. Assuming a car cockpit in which the windows are replaced by indirect vision, we compared the actual driver's eye point with two extremes: a high, third-person viewpoint providing an overview of the surroundings without near blindspots from the body of the own vehicle and a low, first-person viewpoint that improves the visibility of nearby obstacles and increases the amount of optical flow from self motion.

A third study was aimed to provide evidence for usefulness of small continuous changes in viewpoint to replicate motion parallax from head movements and the effect on depth perception, driving performance and experience.

## 7.1 Distortion

Using a novel image transformation that combines a rectilinear center for far distance and compressed side areas for perception of near neighboring lanes, we have been able to provide a field of view of 180 degrees, which essentially eliminates the rear blindspots of mirrors, simultaneously with usable distance perception and minimal, distracting distortion of optical flow. This field of view is significantly wider than that provided by conventional solutions which typically range between about 10 and 30 degrees. Furthermore, this solution provides the significant merits of being able to



---

overview the rear in one glance or the ability to judge not just presence but also time-to-contact. [21]

While the concept of combining an distortion-free region with a distorted region is similar to aspherical door mirrors and systems like FisheyeVision [34], the described method goes further than both by defining and preserving specific properties of the output image that are useful for the given task, like straightness of road boundary or minimum display size of objects. The method also addresses the need for novel, task specific mappings as identified in the FlyViz paper [33] as future research areas and succeeds by proposing a method that combines a wide field of view where useful with far distance perception where needed.

An experimental car was equipped with a prototype implementation that utilizes a high-resolution fisheye camera with internal DSP. An experiment showed that distance perception using our proposed system is on average similar to the standard driver-side door mirror of our experiment vehicle. This result shows that using indirect vision systems with task-specific projections, it is possible to combine a wide field of view and correct distance perception which in conventional, common knowledge have been trade-offs. It should be noted that the proposed solution covers 180 degrees field of view, which is a huge step from the field of view of conventional rear-view mirror solutions that cover less than 30 degrees field of view. Accurate distance judgments are necessary for drivers to safely change lanes and merge.

Participants with experience of using conventional rearview cameras were more exact in that condition compared with first-time users. This hints that lack of familiarity with camera images in general or with utilizing 2D camera and displays could be one reason for the larger variation in distance judgment for the other participants.

When using our prototype or the door mirror, there was a slight overestimation of distances instead of the slight underestimation that would be expected if magnification was the only contributor to distance perception. Interviews of the subjects after all trials had finished revealed that this effect might have been caused by participants incorporating their lack of confidence into their distance judgments and not by other effects from framing or from the lack of binocular depth cues.

Although the “more natural look” and the “less distracting character” of the proposed image when compared with conventional wide rearview images was apparent to most who experienced the prototype and positively commented on, we have not been able yet to evaluate these aspects scientifically.

---

In large, we believe that we have succeeded in manipulating distortion to improve situation dependent spatial awareness. Some issues such as improving the camera's dynamic range (for use at night) and resolution, considering optical flow in curves and when pitching, continuing with further human factors evaluations and adapting to forward vision remain as future work. Another future extension could extend the field of view beyond 180 degrees, resulting for the first time in a first-person view that covers more than 180 degrees but appears natural under self motion.

The method could also be applied outside of driving.

### 7.1.1 Summary

The results show that the projection method can be intentionally manipulated to optimize spatial awareness. In particular, a method was proposed that maps distance in three-dimensional space proportionally to two-dimensional distance on the display and shown to make motion and distance easier to judge compared with conventional projections. The proposed method could be extended to even larger fields of view, views in the direction of motion and to tasks other than driving.

## 7.2 Viewpoints

To answer the question whether manipulating viewpoint can improve spatial awareness, a driving simulator experiment investigated the effect of manipulating viewpoints on driving performance and experience?

### 7.2.1 Simulator experiment

Our hypothesis was that a higher viewpoint would achieve higher situation awareness and enjoyment at a reduced workload, whereas a lower viewpoint would provide stronger motion cues and a more thrilling driving experience. We expected that semantic enhancements would help drivers to make better judgments faster, and that those would add-up when combined with Geometric enhancement. We expected higher viewpoint and AR to be in favour with all participants, whereas a low viewpoint would only resonate with thrill seeking participants.

The enquete showed that in all experiment sections except the static intersection, participants' self-rating about situation awareness, workload and enjoyment were generally higher for the high viewpoint condition compared with the normal viewpoint, whereas the low viewpoint scored about equal to the normal viewpoint. (Figure 5.18) This was also reflected in the mostly positive verbal comments. (Table B.1) These results are not too far from what we had expected, although we suspect that the subjective opinion of

---

participants did not differentiate between the specific aspects asked and should be interpreted with care. For the static intersection, the HMD might not have had enough resolution to present the traffic signs, leading to low scores for the high viewpoint condition.

Task duration is more difficult to interpret. A shorter duration in any of the course sections might be caused by higher situation awareness and less workload if other aspects like correctness of decisions and precision of driving are equal. But a longer duration is not necessarily negative, e.g. if it is caused by higher situation awareness possibly leading to more careful driving. Participants often seemed to simply give up checking safety and start driving blindly when having to avoid close obstacles in the normal and low viewpoint conditions. Less variation in duration could generally be considered positive indicating less difference between drivers and possibility for smoother traffic flow. In some cases, larger variation could be positive if it enables individuals to fully utilize their personal potential for enjoyment by being different from the average. The results (Figure 5.6) hint that high viewpoint has potential to reduce driving duration in some conditions that require surround situation awareness like the dynamic intersection and parallel parking conditions in this experiment. The results also indicate that a higher viewpoint may reduce variation between drivers. We suspect that the high viewpoint might have reduced individual differences in coping with difficulties like blind spots and difficult driving maneuvers like parallel parking. The high viewpoint might also have motivated participants to drive in a way that they thought was more acceptable when watched from a third person perspective and that might have led to less variation between participants. Given that neither very short nor very long task durations seem ideal, there might be some ideal range for the duration of any given task. That range might differ between participants, at least in the case of viewpoints where different skill levels affect driving performance, though possibly less for objective viewpoints.

In the dynamic intersection task where participants had to watch cross traffic from left and right and judge their speed and safe distance before crossing the intersection, a task that may occur in everyday driving, participants were able to make more correct decisions faster with less range of head motion when using the high viewpoint whereas the low viewpoint was similar to normal viewpoint. The minimum distance between cross traffic and participants increased slightly for high and low viewpoints, indicating that participants were either better at judging safe distances or inclined to be more careful in their driving behavior.

---

In the static intersection task where participants had to understand traffic signs, notice the presence of obstacles like pedestrians and parked cars, make a judgment about the direction to take and execute it by actually driving into the chosen direction, which also can be considered a basic task in everyday driving, the high viewpoint enabled participants to drive with less range of head motion and likely less physical effort but had no effect on duration or correctness of decisions. Some participants mentioned that the resolution of the HMD was too low limiting the readability of the traffic signs.

The curve and crank conditions which we initially assumed to give similar results led to results that differed from each other. In the curve section, high viewpoint was not a benefit with regard to average distance driven off-road and even increased the variation between participants, while low viewpoint doubled the average distance driven off-road inspite of better visibility of the poles marking the road boundary on both sides. These results could be interpreted as the high viewpoint lacking in optical flow required for curve negotiation, thereby cancelling the benefit of overview, but also lack of being used to this view for driving. Some participants were caught verbally reporting that they like the overview in the high viewpoint condition that enables them to see literally “everything” but that somehow, they could not avoid veering off-road. The low viewpoint might have reduced overview of the road trajectory and made it more difficult to steer. In contrast with the high viewpoint, the low viewpoint increased optical flow from the road and the poles that led to a higher perceived speed which might have affected the ability for smooth curve negotiation. The crank was a different beast entirely as it required only straight driving interrupted by 90 degree turns at the right moment. Here, the high viewpoint was beneficial with respect to distance driven off-road compared with normal viewpoint for 4 participants while equal and worse for 1 participant each. The low viewpoint was only beneficial for 2 and equal for 4. This might be indicative of both overview and detail helping with timing decision based on the surrounding situation though through different mechanisms and to different degrees.

A smooth parallel parking maneuver required 30 to 40 seconds. (Figure 5.24) While high viewpoint resulted in durations either within or close to that range, in the normal and low viewpoint conditions it happened in 4 (or 2 each) out of 12 maneuvers that participants would fail to find the smooth trajectory in the first attempt and end up with a lot of corrections that took more than 60 seconds. This is also confirmed by less direction changes in the high viewpoint condition.

For the parking exit task, duration is similar between normal and high viewpoint although the number of direction changes increased for 3 participants.

---

This was likely due to the overview providing more information about distance to nearby obstacles thereby motivating some of the participants to add an additional reversing maneuver. The duration increased significantly for the low viewpoint condition. This might have been caused unintentionally by the life-size, immersive rearview that seemed to confuse some of the participants.

All in all, the high viewpoint seemed to be beneficial in situations that require overview over the surrounding situation in subjective ratings of situation awareness, anxiety, fun to drive and mental and physical effort as well as in objective performance measures. While HRV data was inconclusive, GSR data hinted that the high viewpoint on average reduced mental stress in the dynamic intersection and parallel parking conditions. High viewpoint may not be superior in tasks like curve negotiation in which the driver needs to be immersed in the environment to receive cues from optical flow though it was generally at least on par with the normal viewpoint. We did not test whether situations that require immediate, unconscious reaction like emergency avoidance maneuvers in response to imminent dangers (without requiring an overview) might be difficult to handle in the high, third person viewpoint, or whether it may be easier to get distracted in third person viewpoint.

We were not able to identify clear benefits of the lower viewpoint in our experiment. One limitation could have been that a more significant reduction in eye point height, which would have increased the thrill of driving by increased optical flow from the closer ground surface and surroundings that are relatively higher when compared with the eye point, was not possible using the current setup because of the tendency to induce stronger motion sickness. The low viewpoint condition was still in the range of eyepoint height of a sports car. While there were no clear, significant trends, some individuals seemed to respond and drive differently from the normal viewpoint condition. It may therefore be worth studying differences in driving with respect to eyepoint height in a real environment using real cars.

While the viewpoint was fixed or automatically chosen in the experiment conditions, participants could control the viewpoint height by their head pose during the initial practice run. Though the current implementation may have been too sensitive to head motion causing slight motion sickness in some of the participants, giving drivers the ability to actively choose their viewpoint is an intriguing possibility, as it may not be always possible to automatically choose a single, best, ideal viewpoint. It could also avoid feeling having to commit your fate to the automatic viewpoint selection. Actively changing the viewpoint beyond the flexibility of the human body is

---

something that requires time getting used to, but should be possible given that car drivers often unconsciously use head motion to widen the field of view in their rear mirrors. Participants will need more time to get used to such a system. Gestures could be an alternative to controlling by head pose.

From our observations in the driving task, it would seem that a high viewpoint is preferable over normal viewpoint also when informing passengers of an autonomous, driverless vehicle about the surrounding situation if that need should occur, as the overview makes it easier to understand the whole situation in shorter time without continuous watching.

### 7.2.2 Interaction of Changes in Viewpoint with Augmented Reality

Augmented Reality (AR) was implemented and compared in the static intersection condition which required awareness of traffic signs and obstacles and in the dynamic intersection condition which required awareness and judgment of speed and safe distances of cross traffic from left and right directions.

While the average task duration for each section decreased slightly in both, and for 4 participants in the former and 5 in the latter condition, the details seem more complicated.

Verbal comments indicated that the virtual safe distance indicator in the dynamic intersection was neither completely understood during the short experiment time nor completely trusted. While it was rendered semitransparent to distinguish it from the “real” objects in the environment, some participants still feared getting in contact with the virtual augmentation, which might have led to longer missed opportunities for driver viewpoint. (Figure 5.10) The same figure indicates that AR was effective for the low viewpoint, which could have been due to being less accustomed to that eye-point height and wanting any help one could get, or that it actually did help in that case because approaching speed was more difficult to judge from a low viewpoint with less perspective.

For high viewpoint, there was no additional benefit of reducing waiting time as the perspective overview alone exhausted the improvement potential.

When AR was trusted and relied upon, it actually seemed to reduce the minimum distance to cross traffic by more risky maneuvers as there was no motivation to add additional safety margins. A more sophisticated gradual safety indicator might alleviate this issue, but this nonetheless hints at the difficulty of designing augmentation.

The detailed analysis of enquete results in subsection 5.3.3.4 reveals that there seem to be some significant differences in how people think of and

---

utilize AR. Half of the participants indicated no clear difference in self-ratings for AR vs no-AR. One participant thought of it as entertainment. Another participant rated it as beneficial for normal and low viewpoint, but not for the high viewpoint. Only one participant rated it as beneficial for all viewpoints. In most cases, if it was rated positively, it was often rated as being positive in all aspects, which could either be true or be showing that self-ratings tend not to differentiate details.

From these results, it seems that AR might appeal only to some users and that it requires careful design as it needs to be trusted and have as little negative side effects as possible. It should be noted though that for example the blind spot warning systems that have become rapidly common in recent years and can be considered a simple form of AR have quickly become indispensable after several years of lingering at low market penetration when it didn't appeal to most.

From these observations of driving, it seems to us that Semantic enhancement would be beneficial for informing drivers of a shared-driving car and passengers of an autonomous car about the surrounding situation, enabling them to disregard less important objects and quickly understand the decisions of the intelligent car system.

It is not clear how Geometric and Semantic enhancement should be designed relative to each other, especially in cases where they address the same issues. Should they be clearly separated, possibly giving priority to one over the other for a given set of limitations, or is it okay to mix them up? And if both enhancements address the same issue, should they be designed so that their effects add up or does that not matter?

The lack of effect of AR for high viewpoint in Figure 5.10 might be a case where the Geometric enhancement alone in high viewpoint without AR and the Semantic enhancement in low viewpoint with AR both improved so much that the additional combined enhancement in high viewpoint with AR had no further effect. The subjective rating of participant 4 in subsection 5.3.3.4 also seems to share this view. It may actually be natural to assume that there is no additional benefit of combining Geometric and Semantic enhancement if both address the same issues and the effect of each don't add up.

Although we were not able to implement and confirm cases where these two enhancements did add up, we can think of cases where for example details might be added to a perspective overview by AR or information about the surroundings could be added by AR to a perspective detail view.

---

The lack of effect of combining viewpoint changes with AR in this experiment might have been because each alone improved so much and in the same aspect that there was no additional combined enhancement to benefit from.

### 7.2.1 Prototype

The aim of this implementation was to show that viewpoints can be manipulated in a real, automotive cockpit using current technologies. A prototype capable of showing the surroundings from any viewpoint was successfully implemented. While we could not obtain permission for experimental evaluation within our facility, we collected feedback from test drives that included both positive and negative opinions.

Situational awareness was rated high for both the birdview as well as for the life-size rear view. Some participants felt that properly utilizing the birdview would require some time getting used to, and that the automated selection of point of view was sometimes confusing and different from what they expected. In contrast, the life-size rearview was immediately useful, but abrupt switching between forward and rearward view, though triggered by driver making a gear change, felt unnatural to some. The rotation direction of the reversed image opposite to the rotation of the car and the reversed movement direction caused motion sickness in some. It is unclear which of these were stronger; the rotational direction could be corrected with matched rotation to the steering operation in a car with a steer-by-wire steering system.

Concerning AR, the unnatural environment of the roads within our facility and the unnatural character of a test drive limited both possible contents and the resulting experience. The comments hint that the implemented items were at least understood. Effectiveness and acceptance of AR visualizations need to be investigated more deeply and in a more realistic environment to draw any conclusion.

### 7.2.2 Summary

- A high viewpoint seems to improve spatial awareness and resulted in high subjective ratings by all participants. Nevertheless, a high viewpoint is not always the best viewpoint, for example when negotiating a curve.
- Results from the low viewpoint that was implemented as comparison were similar to the actual eye point. A larger difference between these viewpoints might have resulted in measurable differences but was not used due to motion sickness.



- 
- AR helped some but not all participants to make correct judgments faster. While the high viewpoint was rated highly by all participants, there were large differences in how participants rated AR. The reason for these differences could stem from differences in understanding and trust of the presented AR. Changes in viewpoint and AR are combinable and should add-up in theory, but we were not able to identify such cases in our experiment.
  - We successfully implemented a prototype capable of manipulating the viewpoint by constructing a 3D model of the surroundings using omnidirectional cameras and laser range finders. Participants were interviewed after test rides, revealing a general positive stance but also a lot of technical and conceptual issues that need to be addressed.

### 7.3 Motion Parallax from Head Motion

We tested the contribution of binocular stereo and motion stereo to driving, as an essential design question separate from issues like resolution or dynamic range which have and can be expected to continually improve over time. Both are related with projection as they are small differences or changes in viewpoint. Our expectation prior to the experiment were that driving paths would show a large variation in the monocular base case and that adding binocular and further motion cues to a monocular base condition would gradually reduce that variation due to participants becoming able to make more accurate distance judgments. While binocular stereo improved driving over monocular condition, further adding motion stereo did not. We did not test a motion stereo only condition and cannot make final judgments but believe from our experience using the system that motion stereo was not very effective given the small range of motion during driving and the presence of other, strong cues for depth like optical flow. Our results match the results described in [5] that failed to show differences in driving performance resulting from motion stereo in a driving simulator. While we believe that our results would also hold for higher resolution images than in our implementation, this has not been proven.

Unexpectedly, we found that the distance between camera position and the actual head influenced the path of the car when making tight turns profoundly. This seems to be a strong effect that needs to be considered when designing indirect vision cockpits. Interestingly, we did not observe this issue in the other prototypes:

- For the rearward vision system described in Chapter 4, while the offset between the eye point and the camera location in the rear is large, the image is likely not used for deciding the accurate trajectory. We

---

might have been able to measure the effect if we had tested a distance of 0m between the rear of the own car and the front of the target car even though the current experiment setup would not have allowed us to place the reference car at 0m in front of the eye point.

- The prototype described in Chapter 5 had a mode similar to this prototype but a curved two display and camera layout with less longitudinal offset for the display on the left. The prototype also had other modes including first-person viewpoints from the actual eye position and third-person viewpoints which are free from this issue.

It might be enough to have parts of the own car visible in the camera images to achieve an accurate awareness of the viewpoint or the location of surrounding objects relative to the car.

### 7.3.1 Summary

- We implemented an indirect vision cockpit that replicates the continuous motion parallax of direct vision using a robotic camera that tracks the head motions of the driver. Due to engine vibrations causing the camera and the captured image to vibrate the results may contain artifacts. We measured the path of four participants driving through an obstacle avoidance course but were not able to identify trends. The results show that motion parallax from head motion does not seem to measurably improve driving performance even though it may improve the experience especially at standstill and slow speed.
- The offset between the head position and the camera position may lead to an offset in perceived location and needs to be taken care of when designing indirect vision cockpits.

---

## 8. Conclusion and Future Work

This thesis investigated the potential of indirect vision systems to improve human performance in visual tasks. A unique characteristic of electronic indirect vision systems is their capability to freely manipulate the projection method. We therefore looked at the possibility of manipulating the projection method to enhance spatial awareness. We focused on car driving as a popular task that requires a high level of spatial awareness and allows the system to be implemented as part of the car cockpit without the user having to carry the system including sensors, image processing system and displays. Having chosen car driving as the example application does neither mean that the methods and results are specific to cars and car driving nor that they are limited to manual driving and not applicable to possible future driverless cars. If anything, we expect driverless cars to be more likely to be equipped with indirect vision systems than manually driven cars, given that there is no driver who needs to watch the surroundings and be in control of the car, the possibility of using the displays for other purposes like entertainment or work, and a higher demand for privacy while doing tasks other than driving.

### 8.1 Major results

- We have implemented an indirect vision system for cars aimed to replace direct vision through windows. By combining multiple cameras and laser range finders to create a 3D model of the surroundings in real time, we were able to display the surroundings in three dimensions from any viewpoint as a computer graphics representation textured by the camera images. This system is the so far most powerful, publicly known such implementation.
- A simulator experiment implementing a sequence of typical driving tasks showed that manipulating the point of view depending on the driving situation may improve both driving performance and experience, suggesting the future potential of indirect vision applications.
- A novel projection method was proposed that combines properties from rectilinear perspective projections and parallel projection to map longitudinal distance in space proportionally to two-dimensional distance in screen coordinates without distorting the direction of optical flow from self motion. A prototype implementation covered a field of view of 180 degrees on a small display, thereby far exceeding the typical field of view of 20 to 30 degrees provided by conventional solutions and refutes the

---

common wisdom that correct distance judgments cannot be achieved in images with extremely wide field-of-view.

- The proposed method can eliminate the need to scan multiple visual locations to achieve awareness of the whole situation which can be useful in driving and other tasks that require spatial awareness of the surroundings. It is also likely to improve time-to-contact predictions.
- We learned that the offset between camera location and eye point may influence trajectory and needs to be considered.

## 8.2 Future work

Sensor, image processing and display technology need to improve in resolution, dynamic range, sensitivity, delay, ghosting, etc before indirect vision will really match and exceed the experience of natural direct vision. The system will also have to be designed with a backup in case of failure, either in the image pipeline or with an autonomous system that can take over control from the driver when needed adding to the cost of the system. On a different note, any such system should be resilient to hacking attempts.

Concerning the use of third person viewpoints, our results and experiences hint that the following issues deserve further study: Ability to maneuver through curves in third person viewpoints that are likely to have less optical flow cues compared with immersive first-person viewpoints, ability to react to imminent danger without the optical flow directed towards the viewer typically present in first person viewpoints, and susceptibility to distraction and other mid and long-term effects.

Concerning the use of variable viewpoints, we experienced that a selection by the system might not always be optimal in a given situation or might not appropriately consider the user's intention. A method should be developed that allows the user to intuitively select a desired viewpoint.

Manipulating the point of view may be useful for tasks other than driving where spatial awareness is useful. Smaller sensors and displays will make it possible to implement a wearable version.

The concept for field of view expansion described in section 4 should be extendable to forward vision. A near area in which distances are mapped linearly may not feel natural as human perception might be more tuned to optical flow during forward than rearward motion, therefore a slightly accelerating optical flow might be desirable. The concept could also be extended to larger field of view, resulting in a image with 200+ degrees of forward field of view on a 20+ inch size display. Another extension should be dynamic

adaptation to driving through curves. During curve driving, the optical flow is curved, and the curvature of the optical flow should be preserved.

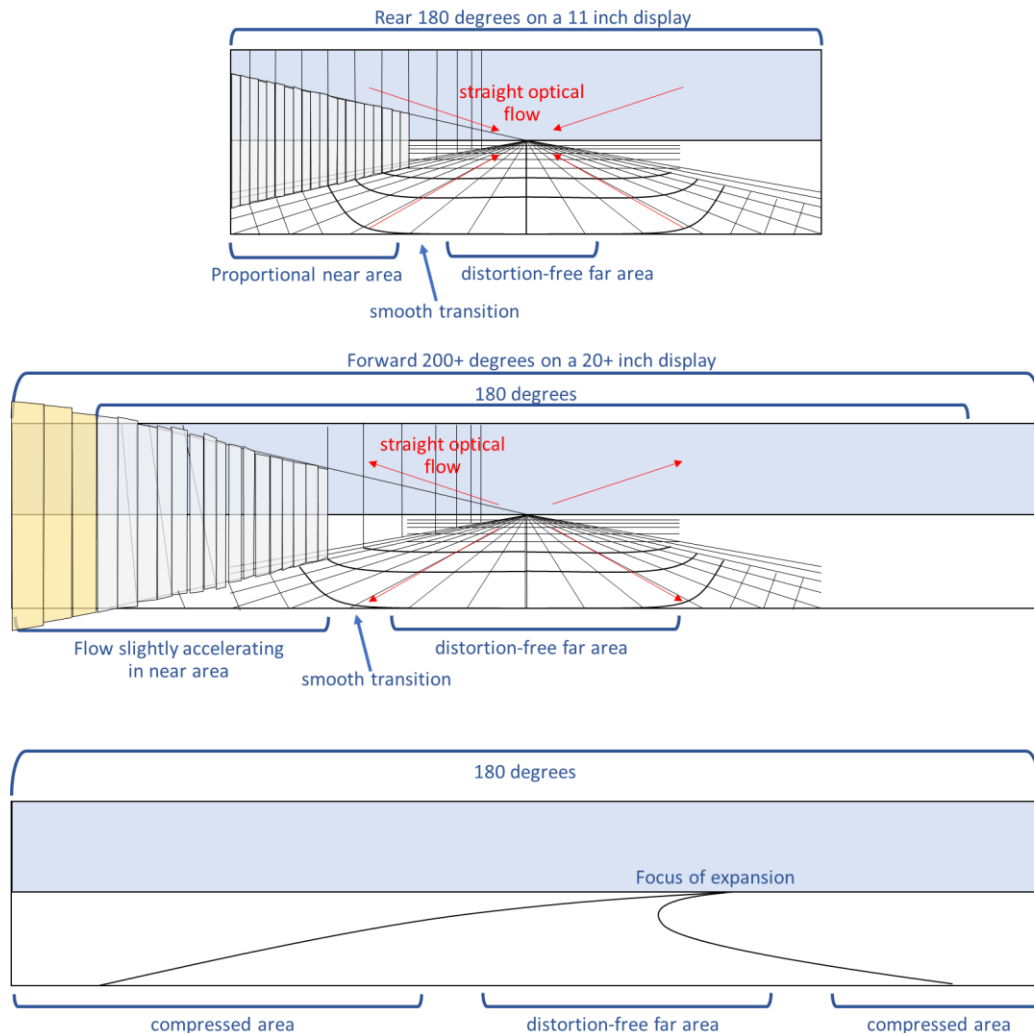


Figure 8.1: Outline of future possible extensions of the method to expand field of view

Motion sickness caused by the disagreement of visually perceived movement and the vestibular system's sense of movement is an issue that needs to be solved in the future. The differences can be temporal, caused by the delay in capturing, processing and displaying the image, or geometric, caused by unintended or intended distortions, lack or inaccuracies in motion parallax, the close distance to a flat display with wrong oculomotor cues, and others. By determining the maximal acceptable amount of disagreement and engineering the indirect vision system to satisfy those thresholds, motion sickness is in principle solvable. Considering that humans can get used to using

---

prescription glasses that distort the viewing field without getting motion sick, solving motion sickness from distortion is not necessarily unsolvable.

The future work described would make an intermediate indirect vision cockpit possible in the near future in which forward and rearward views are displayed on separate monitors of about 20 and 10-inch diagonal size respectively providing a complete surround view within the comfortable field of view of human perception. Some driver pose recognition may be used to control changes of point of view by driver intention and small windows may be used as a fallback mechanism in rare situations of system failure. A final indirect vision cockpit for a driverless car may display the surroundings as small picture-in-picture regions within a large surrounding curved display covering the whole interior.

Wearable solutions could switch between different views depending on the situation and utilize the field of view expansion method described while moving.

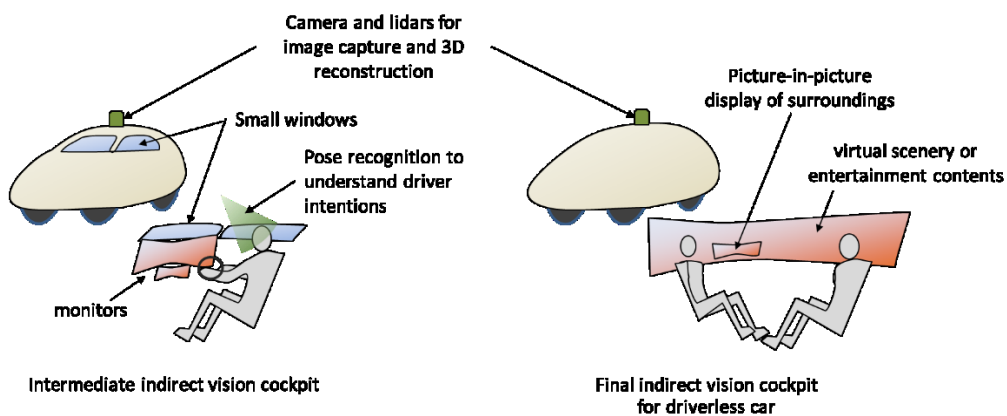


Figure 8.2: Possible development of future indirect vision cars

---

## Bibliography

- [1] "Automotive News Europe," [Online]. Available: <http://europe.autonews.com/>.
- [2] "Vienna Convention on Road Traffic," 1968. [Online]. Available: <http://www.unece.org/fileadmin/DAM/trans/conventn/crt1968e.pdf>.
- [3] R. H. A. Cohen, "Zur Bedeutung des fovealen Sehens fuer die Informationsaufnahme bei hoher Beanspruchung," *Sicht und Sicherheit im Strassenverkehr*, Verlag TUV Rheinland, Koeln, 1990.
- [4] J. Cutting and P. Vishton, "Perceiving Layout and Knowing Distance: The Integration," in *Perception of Space and Motion*, W. Epstein & S. Rogers (Eds.), New York: Academic Press, 1995, pp. 69-118.
- [5] D. Stewart, C. J. Cudworth and J. R. Lishman, "Misperception of time-to-collision by drivers in pedestrian accidents," *Perception*, vol. 22, pp. 1227-1244, 1993.
- [6] J. A. Hultgren, B. Blissing and J. Jansson, "Effects of motion parallax in driving simulators," in *Driving Simulation Conference*, Paris, France, 2012.
- [7] J. D. Pfautz, "Depth perception in computer graphics," *Technical Report*, University of Cambridge, UCAM-CL-TR-546, 2002.
- [8] N. Benguigui, H. Ripoll and M. P. Broderick, "Time-to-Contact Estimation of Accelerated Stimuli Is Based on First-Order Information," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 29, no. 6, p. 1083-1101, 2003.
- [9] J. López-Moliner, A. Maiche and S. Estaún, "Perception of acceleration in motion-in-depth with only monocular and both monocular and binocular information," *Psicológica*, vol. 24, pp. 93-108, 2003.
- [10] S. Durkee and N. Ward, "Effect of Driving Simulation Parameters Related to Ego-Motion on Speed Perception," in *PROCEEDINGS of the Sixth International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design*, 2011.
- [11] F. Panerai, J. Droulez, J.-M. Kelada, A. Kemeny, E. Balligand and B. Favre, "Speed and safety distance control in truck driving: comparison of simulation and real-world environment," in *DSC 2001 Driving Simulator Conference*, Sophia Antipolis, September 2001.

- 
- [12] J. Duijnhouwer, R. J. A. v. Wezel and A. V. v. d. Berg, "The role of motion capture in an illusory transformation of optic flow fields," *Journal of Vision*, vol. 8, no. 27, April 2008.
- [13] M. R. Endsley, "Toward a theory of situation awareness in dynamic systems.," *Human Factors*, Vol.37(1), pp. 32-64, 1995.
- [14] J.Cegarra et al., "A neuroergonomics perspective on mental workload predictions in Jens Rasmussen's SRK framework," *Le travail humain* 80(1):7, April 2017.
- [15] M. P. Reed, M. M. Lehto and M. J. Flannagan, "FIELD OF VIEW IN PASSENGER CAR MIRRORS," UMTRI Technical Report, UMTRI-2000-23, June 2000.
- [16] H. Hecht and J. Brauer, "Convex Rear View Mirrors Compromise Distance and Time-To-Contact Judgements," *Ergonomics*, vol. 5, no. 4, pp. 601-604, 2007.
- [17] M. Flannagan, M.Sivak, J. Schumann, S. Kojima and E. Traube, "Distance Perception in Driver-Side and Passenger-Side Convex Rearview Mirrors: Objects in Mirror are more complicated than they appear," *UMTRI Technical Report UMTRI-97-32, Ann Arbor, Michigan: University of Michigan Transportation Research Institute*, 1997.
- [18] A. D. Vos, "Non-Planar Drivers Side Rearview Mirrors: A Survey of Mirror Types and European Driver Experience and a Driver Behavior Study on the Influence of Experience and Driver Age on Gap Acceptance and Vehicle Detection," *Technical Report DOT HS 809 149, Washington, D.C.: National Highway Traffic Safety Administration, U.S. Department of Transportation*, 2007.
- [19] A. R. Perline, A. Hicks and K.Ronald, "Blind-spot Problem for Motor Vehicles," *Applied Optics*, vol. 44, pp. 3893-3897, 2005.
- [20] "Ford's Blind Spot Mirror," 8 8 2008. [Online]. Available: [www.autosavant.com/2008/08/08/fords-blind-spot-mirror/](http://www.autosavant.com/2008/08/08/fords-blind-spot-mirror/).
- [21] Nissan Motor Corporation, "Around-view Monitor System," [Online]. Available: <http://www.nissan-global.com/JP/TECHNOLOGY/OVERVIEW/avm.html>.
- [22] M. Flannagan and M. Sivak, "Driver Workload for Rear-Vision Systems with Single versus Multiple Display Locations," *SAE Technical Paper Series, 2005-01-0445*, 2005.
- [23] M. Flannagan, M. Sivak, J. Simpson, "The Role of Binocular Information for Distance Perception in Rear-Vision Systems," *SAE Technical Paper Series, 2001-01-0322*, 2002.



- 
- [24] M. Flannagan , M. Sivak, “Framing Effects on Distance Perception in Rear-Vision Displays,” *SAE Technical Paper Series, 2003-01-0298*, 2003.
- [25] C. Smyth, G. Gombash , P. Burcham, “Indirect Vision Driving With Fixed Flat Panel Displays for Near-Unity, Wide and Extended Fields of Camera View,” *Tech. rep., Army Research Laboratory, Aberdeen Proving Ground, MD*, June 2001.
- [26] Nissan Motor Corporation, "Smart Room Mirror," [Online]. Available: [http://www2.nissan.co.jp/OPTIONAL-PARTS/NAVI\\_AUDIO/UPGRADE/SMART\\_ROOMMIRROR/](http://www2.nissan.co.jp/OPTIONAL-PARTS/NAVI_AUDIO/UPGRADE/SMART_ROOMMIRROR/).
- [27] T. L. Brown, Y. He, C. Row and T. Schnell, "Is More Better? — Night Vision Enhancement System’s Pedestrian Warning Modes and Older Drivers," *Annals of advances in automotive medicine* , vol. 54, no. January, pp. 343-50, 2010.
- [28] "GPV Colonel," [Online]. Available: <http://www.army-technology.com/>.
- [29] O. Tsimhoni and P. Green, "Night Vision Enhancement Systems," *Technical Report UMTRI-2002-05*, April 2002.
- [30] T. Yoshida, K. Jo, K. Minamizawa, H. Nii, N. Kawakami and S. Tachi, "Transparent Cockpit: Visual Assistance System for Vehicle Using Retro-reflective Projection Technology," in *IEEE Virtual Reality 2008*, Reno, Nevada, USA.
- [31] "Car and Driver Magazine," [Online]. Available: <http://www.caranddriver.com>.
- [32] "Sony SC-1," [Online]. Available: <https://news.mynavi.jp/article/20171110-5g-future01/>.
- [33] Japan Automotive Research Institute, "JARI-ARV," [Online]. Available: <http://www.jari.or.jp/tabid/290/Default.aspx>.
- [34] J. Ardouin, A. Lecuyer, M. Marchal, C. Riant and E. Marchand, "Flyviz: a novel display device to provide humans with 360 vision by coupling catadioptric camera with hmd," in *Proceedings of the 18th ACM symposium on Virtual reality software and technology*, 2012.
- [35] J. Orlosky, Q. Wu, K. Kiyokawa, H. Takemura and C. Nitschke, "Fisheye Vision: Peripheral Spatial Compression for Improved Field of View in Head Mounted Displays," in *Proceedings of the 2nd ACM Symposium on Spatial User Interaction*, 2014.

- 
- [36] K. Fan, J. Huber, S. Nanayakkara and M. Inami, "Spidervision: Extending the Human Field of View for Augmented Awareness," in *Proceedings of the 5th Augmented Human International Conference*, 2014.
- [37] Y.-T. Lin, Y.-C. Liao, S.-Y. Teng, Y.-J. Chung, L. Chan and B.-Y. Chen, "Outside-In: Visualizing Out-of-Sight Regions-of-Interest in a 360 Video Using Spatial Picture-in-Picture Previews," in *UIST 2017*, Quebec City, QC, Canada, October 22-25, 2017.
- [38] Y. Yano, J. Orlosky, K. Kiyokawa and H. Takemura, "DynamicViewExpansionforImprovingVisualSearch inVideoSee-throughAR," in *International Conference on Artificial Reality and Telexistence Eurographics Symposium on Virtual Environments*, 2016.
- [39] E. Niforatos, A. Fedosov, I. Elhart and M. Langheinrich, "Augmenting Skiers' Peripheral Perception," in *ISWC '17*, MAUI, HAWAII, USA, 2017.
- [40] G. Gorisse, O. Christmann, E. A. Amato and S. Richir, "First- and Third-Person Perspectives in Immersive Virtual Environments: Presence and Performance Analysis of Embodied Users," *Frontiers in Robotics and AI*, vol. 4, July 2017.
- [41] H. G. Debarba, E. Molla, B. Herbelin and R. Boulic, "Characterizing Embodied Interaction in First and Third Person Perspective Viewpoints," in *IEEE Symposium on 3D User Interfaces (3DUI)*, Arles, 2015.
- [42] A. Denisova and P. Cairns, "First Person vs. Third Person Perspective in Digital Games: Do Player Preferences Affect Immersion?," in *CHI 2015*, Seoul, Republic of Korea, April 18-23, 2015.
- [43] T. Kosch, R. Boldt, M. Hoppe, P. Knierim and M. Funk, "Exploring the Optimal Point of View in Third Person Out-of-Body Experiences," in *PETRA '16*, Corfu Island, Greece, June 29-July 01, 2016.
- [44] T. Miyaki and J. Rekimoto, "LiDARMAN: Reprogramming Reality with Egocentric Laser Depth Scanning," in *SIGGRAPH ' 16: ACM SIGGRAPH 2016 Emerging Technologies*.
- [45] M. Thoby, "Photographic lenses projections: computational models, correction, conversion...," 23 11 2012. [Online]. Available: [http://michel.thoby.free.fr/Fisheye\\_history\\_short/Projections/Fisheye\\_projection-models.html](http://michel.thoby.free.fr/Fisheye_history_short/Projections/Fisheye_projection-models.html).
- [46] M. Flannagan , M. Sivak, "Distance Perception in CameraBased Rear-Vision Systems," *SAE Technical Paper Series*, 2002-01-0012, 2002.

- 
- [47] M. Flanagan and M. Sivak, "Distance Cues and Fields of View in Rear Vision Systems," *SAE Technical Paper Series, 2006-010947*, 2006.
- [48] T. Yanagi, N. Shimomura, S. Chinomi, H. Mouri, "A Novel Camera-Based Rear Vision System for Lane Changes with 180 Degree Field of View," *FISITA*, 2010.
- [49] T. Yanagi, N. Shimomura, "An Image Transformation Method that Realizes a Field of View of 180 Degrees without Compromising Distance Perception for a Camera-Based Rear Vision System for Merging and Changing Lanes," *First International Symposium on Future Active Safety Technology toward zero-traffic-accident*, September 5-9, 2011.
- [50] T. Yanagi, C. L. Fernandez, M. Y. Saraiji, K. Minamizawa, S. Tachi, N. Kishi, "Transparent Cockpit Using Telexistence," *IEEE Virtual Reality Conference*, pp. 311-312, 23-27 March, 2015, Arles, France.
- [51] J. S. Albus, "4D/RCS A Reference Model Architecture for Intelligent Unmanned Ground Vehicles," *Proceedings of the SPIE 16th Annual International Symposium on Aerospace/Defense Sensing, Simulation and Controls, Orlando, FL*, April 1-5, 2002.
- [52] J. Chen, R. Oden, C. Kenny, J. Merrit, "Effectiveness of Stereoscopic Displays for Indirect-Vision Driving and Robot Teleoperation," *Tech. rep., Army Research Laboratory, Aberdeen Proving Ground, MD*, August 2010.
- [53] C. L. Fernando, M. Furukawa, T. Kurogi, S. Kamuro, K. Sato, K. Minamizawa, S. Tachi, "Design of TELESAR V for Transferring Bodily Consciousness in Telexistence," *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems, Vilamoura, Algarve, Portugal*, pp. 5112-5118, Oct. 2012.
- [54] S. Mahlke, D. Roesler, K. Seifert, J. Krems, M. Thuring, "Evaluation of six night vision enhancement systems: Qualitative and quantitative support for intelligent image processing," *Human Factors*, 49, pp. 518-531, 2007.
- [55] K. Yamada, Y. Izawa, Y. Kaneshima, T. Jitsuhiro, S. Kato, K. Takeda, Y. Ishiguro, "Real oriented Virtuality : Immersive Game with Real oriented Virtuality," *IPSI Interaction*, pp. 800-803, 2016.
- [56] Pioneer Corporation, "Pioneer Introduces the New CYBER NAVI Car Navigation system in Japan With the world's first Head-Up Display to project augmented reality information in front of the windscreen (News Release)," May 2012. [Online]. Available: <http://pioneer.jp/en/news/press/2012/pdf/0511-1.pdf>.

- 
- [57] S. Tachi, K. Minamizawa, M. Furukawa , C. L. Fernando, “Telexistence - from 1980 to 2012,” *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems, Vilamoura, Algarve, Portugal*, pp. 5440-5441, Oct. 2012.
- [58] S. Nagata, "How to Reinforce Perception of Depth in Single Two-Dimensional Pictures," in *S.Ellis (Ed.), Pictorial Communication in Virtual and Real Environments*, London: Taylor & Francis., 1993, pp. 527-545.
- [59] T. Litman, "Autonomous Vehicle Implementation Predictions, Implications for Transport Planning," *Victoria Transport Policy Institute*, May 1, 2017.
- [60] H. Rosenfield, "Self-Driving Vehicles, the Threat to Consumers," June 13, 2017. [Online]. Available: <http://www.consumerwatchdog.com/>.
- [61] E. N. Marieb , K. Hoehn, *Human Anatomy & Physiology*, Pearson.

---

# Appendix A.

## List of publications

T. Yanagi, N. Shimomura, S. Chinomi and H. Mouri, "A Novel Camera-Based Rear Vision System for Lane Changes with 180 Degree Field of View," *FISITA*, 2010.

T. Yanagi and N. Shimomura, "An Image Transformation Method that Realizes a Field of View of 180 Degrees without Compromising Distance Perception for a Camera-Based Rear Vision System for Merging and Changing Lanes," *First International Symposium on Future Active Safety Technology toward zero-traffic-accident*, September 5-9, 2011.

T. Yanagi, C. L. Fernandez, M. Y. Saraiji, K. Minamizawa, S. Tachi and N. Kishi, "Transparent Cockpit Using Telexistence," *IEEE Virtual Reality Conference*, pp. 311-312, 23-27 March, 2015, Arles, France.

# Appendix B. Implementation Details of the System in Chapter 4.

## B.1 Numerical approximation of the desired projection

The steps for computing a numerical approximation to the desired projection is described in the figure below.

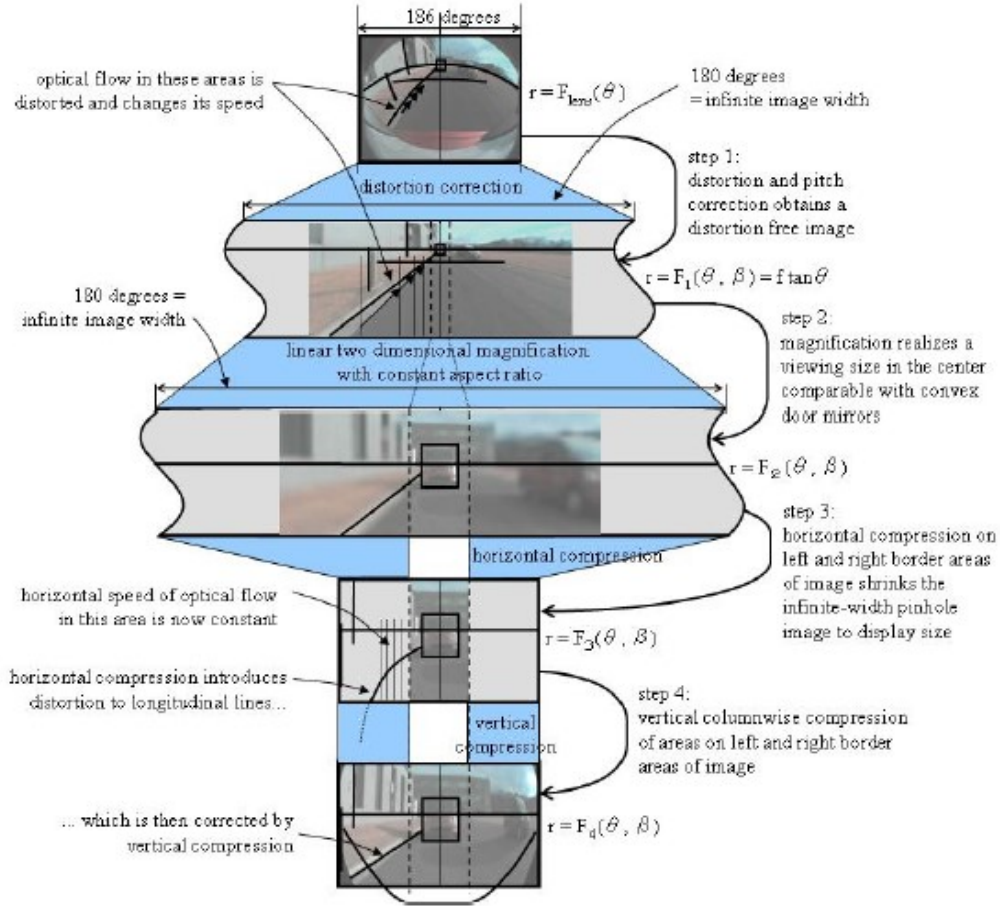


Figure B.3: Step-by-step explanation of our image transformation

The numerical approximation starts with the representation of the perspective, rectilinear projection

$$r = F_1(\theta, \beta) = f \tan \theta \quad (0^\circ \leq \theta < 90^\circ \text{ and } 0^\circ \leq \beta < 360^\circ) \quad (3)$$

where  $f$  is the focal length. This representation is then refined in successive steps, first, to a projection model

$$r = F_2(\theta, \beta) \quad (4)$$

that includes a linear magnification to achieve the wanted magnification in the center part of the image, then into

$$r = F_3(\theta, \beta) \quad (5)$$

that includes a progressive, row-wise horizontal compression in left and right border areas that normalizes the speed of horizontal optical flow, and the final projection model

$$r = F_4(\theta, \beta) \quad (6)$$

that includes a column-wise vertical compression to correct distortion of longitudinal lines introduced by the horizontal compression in the step before.

The projection model  $r = F_{\text{lens}}(\theta)$  of the fisheye camera was obtained from the manufacturer of the camera ignoring individual build tolerances. Alternatively, it may be obtained by a calibration of the internal parameters using known calibration methods. It is then easy to calculate the pixel-wise correspondence between the fisheye image and the wanted projection model  $r = F_4(\theta, \beta)$  and process the image transformation from fisheye image to the proposed projection in one, direct transformation step.

## B.2 Implementation details

Our transformation magnifies the rear center area relative to the areas on left and right side. The fisheye image should therefore have higher resolution than the transformed image. We first recorded high resolution images using a full-HD camera and applied pixel-wise exact image transformation on a PC to verify that the transformed image satisfies the wanted properties. Then we implemented the image transformation as a real-time mesh-wise texture-mapping operation using a small camera for automotive applications with a build-in DSP. That camera was then used in the following evaluation experiment.

The full-HD implementation used the lens-camera combination described in the following table to record high resolution fisheye images as .avi movie files to a memory card, and then applied the image transformation to each frame using Matlab on a Personal Computer.

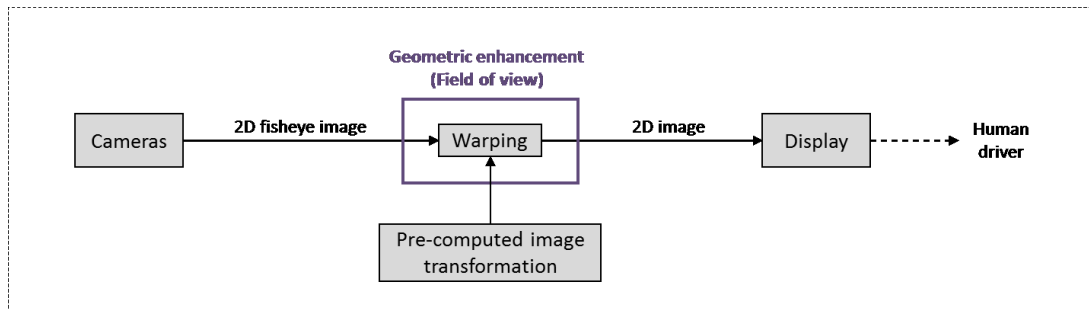


Figure B.4: System architecture

Table B.1 : Specification of the full-HD camera

camera model	Toshiba IK-HR1H
sensor type	CCD, progressive scan
lens	fisheye, Optart FJ06-2K, C-mount (185 deg. horizontal field of view)
resolution of fisheye image	1920 x 1080 pixels
camera size with lens	39 x 39 x 91 [mm] (H x W x D)
resolution of transformed image	1280 x 720 pixels
comments	camera output was recorded to a SD memory card using VISK IR-100 image recorder and post-processed on a PC

To reduce the size of the camera and the total cost of the rear-view system, we implemented the proposed image transformation using a camera with build-in DSP. The build-in DSP is capable of image transformations like digital zooming and distortion corrections eliminating the need for an external image processing unit connected to the camera by a high-resolution interface. Though this system has lower resolution than the full-HD version and in-part suffers from wave-like artifacts caused by using mesh-wise texture mapping instead of pixel-wise exact image transformation, the improvement in radial distortions and the change in distance perception compared with the original fisheye image are unchanged. Considering that this system measures about 1/100th in volume and current device cost compared with the full-HD setup, it performed well. This camera was then used for our vehicular prototype.

Table B.2 : Specification of the automotive camera used

manufacturer	Sony Corporation
sensor type	CMOS, progressive scan
lens	fisheye (186 deg. horizontal field of view)
resolution of fisheye image	1.3M pixels
camera size with lens	23 x 23 x 21 [mm] (H x W x D)
resolution of transformed image	composite video (equivalent to 640 x 480 pixels)
others	internal DSP for image transformation (e.g. distortion correction)



Figure B.5: Image of the camera (with a centimeter scale).



---

# Appendix C. Details of the Experiment in Chapter 5

## C.1 Experiment design

A simulator was chosen over real cars as it allows for faster prototyping and evaluation given the complexity of the total system and ensuring safe operation. In order to achieve a surround 360 degrees field of view as the benchmark natural viewing condition, a head-mount display (HMD) was used instead of a driving simulator with stationary display monitors. The Unity game engine was used in combination with an Oculus Rift HMD. A commercially available car model modified to adjust the difficulty of the driving task to appear close to real driving while being simple enough not to require long customization. An Xbox controller was used for controlling the car to provide a neutral interface not requiring visibility of the own hands in the HMD. The interior of the user's car was shown without a CG rendering of the driver and without rotating steering wheel. The drivers side door mirror was the only working mirror as other mirrors can be obstructed by luggage or by passenger.

## C.2 Participants

6 participants (all males, age 24, 27, 30, 39, 40 and 40) with normal or corrected-to-normal vision and experience driving cars regularly (between 1 and 7 days a week) all completed the same three experiment conditions after a practice run. Participants completed the experiment during regular office hours as part of their work and were not otherwise paid. 1 additional male and 2 females indicated motion sickness from using the HMD, aborted the experiment and were not counted as participants.

## C.3 Procedure

Participants started by filling out a questionnaire, doing a practice run of the experiment course with head-pose controlled viewpoint to understand the experiment task and the viewpoints and completed three experiment conditions in randomized order in which the viewpoint was either a normal, driver's seat viewpoint (condition A), a viewpoint that was raised (condition B) by 8m (parallel parking) or 30m (all other driving situations) or lowered (condition C, by 0.4m). Participants were asked to verbally report their thoughts and feelings while driving. Both intersections were repeated a total of four times in each condition, 2 times without augmented reality and 2 times with. For half of the participants, the AR condition came first. Participants rested for 90 seconds before and 60 seconds after each condition to

record stable data from the GSR and HR sensors, and filled the questionnaire each time after a ride. Participants could abort the experiment any time or add additional rest between conditions.

Table C.3 : Experiment order by participant.

Subject	Experiment conditions in order of execution				Subconditions within Intersection scene	
	Practice	1 <sup>st</sup>	2 <sup>nd</sup>	3 <sup>rd</sup>	1 <sup>st</sup>	2 <sup>nd</sup>
1	D	A	B	C		AR
2	D	A	C	B	AR	
3	D	B	A	C		AR
4	D	B	C	A	AR	
5	D	C	A	B		AR
6	D	C	B	A	AR	

Table C.4 : Experiment design.

		No Sensory enhancement (Day-light, fair weather)		
		No Geometric enhancement	Geometric enhancement	
			Drivers point of view	Detail view (low / rear)
Parking Curve Crank	No Semantic enhancement	Base case	Geometric enhancement	
Intersection (slow/fast)				
	Semantic enhancement	Semantic enhancement	Semantic and geometric enhancement	

#### C.4 Measures

We attempted measuring situation awareness and psychological as well as physical effort comprehensively using objective and subjective measures. Situation awareness can be indirectly assessed from indices like eye movements, directly probed by questionnaires or questions, indirectly measured at the decision-making stage from the decisions made or by verbalizing the decision-making process, indirectly measured by the performance in normal

---

or in emergency situations. High situation awareness can be achieved by employing a high-workload surveillance strategy, so collecting data about psychological and physical effort is necessary for interpreting the results.

The measures for situation awareness included indirect measures that assess situation awareness by performance and indirect indices of situation awareness like feeling of safety and physiological measures like HR, as well as direct subjective reports. Direct objective measures like interrupting the driving task to ask specific questions about the surroundings that participants answer from memory without seeing the surrounding were deemed too difficult to realize with HMDs.

Driving performance was assessed from the duration to complete each section of the driving task, the distance spent off-road with some part of the car outside the road boundary during the curve and the cross sections and the number of correct/incorrect judgments at the static intersection. For the dynamic intersection, the number of crashes (physical contact between the participants car and cross traffic) and near-miss situations (the car of the participant being within 1m in front of cross traffic), the duration of missed opportunity windows in which the participant could have safely crossed the intersection and the minimum distance from the cross traffic to the participant crossing the intersection were measured.

As for biophysical measures, heartrate (HR) and heart rate variability (HRV) were collected using a Garmin chest belt sensor connected with a Garmin Edge 520 and recording heart rate at one data point per second as well as all RR intervals into a .fit file and analyzed using Kubios software. HR provides a rough estimate for physical effort and mental arousal, but these are difficult to isolate from each other. Concerning HRV, pNN50, i.e. the proportion of pairs of successive NNs that differ by more than 50 ms divided by total number of NNs, where NN is the time between “normal” beats originating in the sinoatrial node, is known to be higher in rest conditions than with mental tasks and was calculated for each experiment section.



Figure C.6: HR sensor

---

In addition, Galvanic Skin Response (GSR) was measured using NeuLog GSR sensors attached to two fingers of the non-dexterous hand at a sampling rate of 5 samples per second. We found that the calibrated mode of the sensor with a range of 0 to 10 micro Sieverts at a resolution of 10 nano Sieverts would lead to signal clipping in some instances and had to use the non-calibrated mode with a non-specified but wider range and 16-bit ADC resolution. The unit is designated as “arb” as a shorthand for “arbitrary” in this paper in accordance with the manufacturer of the sensor. In order to isolate the phasic skin conductance response (SCR) from the tonic skin conductance level (SCL), a  $\pm 4$  second median was subtracted from the GSR signal and any pair of a onset ( $>0$  arb) and offset ( $<0$  arb) in a adaptation of the recommended protocol in [iMotions GSR guide]. Due to the non-calibrated signal peak amplitude could not be measured. The collected biophysical signals might have been influenced by simulator sickness and motion sickness from the HMD.



Figure C.7: NeuLog GSR sensor

The pose and location of the HMD was recorded as an approximation of head pose as an index of physical effort.

For subjective self rating, participants were asked to fill a questionnaire before asking about their general driving experience and after each experiment condition asking about the particular experiment. The questionnaire asked for subjective degree of surround situation awareness, degree of incorrect driving like leaving the road boundary or making a wrong decision, degree of feeling of anxiety, degree of driving enjoyment, degree of mental effort and degree of physical effort, on a 5 point scale from 2 (very high) to -2 (very low) for each section and AR condition of the experiment (parking exit, curve, crank, static intersection without AR, static intersection with AR, dynamic intersection without

Table C.5 : Overview of objective and subjective measures for situation awareness and effort.

	Situation Awareness			Effort	
	direct	indirect		psychological	physical
		performance	indices		
objective	Answers to questions during interruptions or after experiment (correctness of replies)	Perception and control (e.g. smoothness, lane keeping), Comprehension (e.g. path across intersection), Projection (e.g. timing judgment at intersection)	Psychological (e.g. HR in response to critical event), Physical (e.g. eye/head motion towards potential danger)	Arousal (e.g. HR)	Amount of eye/head motion
subjective	Self rating of S.A. (should match obj. S.A.)	Self rating of performance (should match obj. performance)	Self rating of feeling of safety (should match obj. performance)	Self rating of enjoyment and workload	Self rating of physical effort

Table C.6 : Overview of objective and subjective measures for situation awareness and effort by experiment stage.

	Experiment task	Measures for Situation Awareness						Measures for Effort					
		direct		indirect				Psychological (Arousal, stress)		Physical			
		Obj.	Subj.	performance		indices		Objective	Subj.	Obj.	Subj.		
baseline	2 min rest, eyes closed, hands on controller	Baseline data for HR/HRV/GSR											
1st situation	Parallel parking exit	• self rating	• self rating	• Min. distance	• self rating	• Head motion	• HR, HRV, GSR	• self rating	• Head motion	• self rating	• Head motion		
2nd / 3rd situation	Curve / Crank			• # direct. change							• Margin to curb	• HR	• Head motion
4th situation	Slow Intersection, decide direction from 3 (w/o AR) or 2 (w/AR) options			• time							• Lane departures	• Head motion	• Head motion
5th situation	Intersection with fast cross traffic (AR: THW decision aid)			• Correctness							• time	• Head motion	• Head motion
6th situation	Parallel parking			• Stopped							• Margin to curb	• HR	• Head motion
6th situation	Parallel parking			• Lane departures							• time	• Head motion	• Head motion
baseline	1 min rest, eyes closed, hands on controller	Baseline data for HR/HRV/GSR											

---

## C.5 Time Sequence of Experiment Data

Figure C.8 shows the time sequence of the high viewpoint condition for participant 4 as an example of a smoothly completed experiment condition. All graphs have the sample index of 5 samples per second as the x axis. In this case, the duration of the graph is 1501 samples or about 300 seconds or 5 minutes. The sections of the course are separated by yellow vertical lines and the repetitions of the intersections by gray vertical lines.

The graph at the top shows the speed in m/s on the left scale and the car heading in degrees on the right scale, where 0 is the initial heading, north, positive degrees are towards east and negative degrees are towards west. We see the car leaving the initial parking slot towards east, then passing the curve first going westwards before turning back eastwards. In the crank section, we see a 90-degree left turn followed by a straight section and a 90 degree right turn with the car ending in northward heading with the participants stopping the car twice. Speed was limited in these sections to about 3 m/s to prevent motion sickness. In the static intersection we see the participants making a right turn in repetition #2 and stopping in repetitions #1 and #3. In the dynamic intersection, the participant corrects the stopping location in repetition #2 at around  $x=1000$ . In the parallel parking section, the speed was again limited to and we see the participant making a complex maneuver starting with a left-right steering maneuver before reversing. The second graph shows the direction of the participants head around the vertical axis in degrees relative to the heading of the car. (0 = straight forward, positive = right, negative = left.) Due to the overview from the high viewpoint, there is almost no head rotation during parking exit, curve, crank and parallel parking sections. In the static intersections we see some head rotation, followed by a high amount of head rotation in the dynamic intersection which is somewhat linked to the frequency of cross traffic approaching and passing in front of the participants car and indicated by the distance to cross traffic on the right scale.

The curve depicted in green in the bottom graph shows HR measured in bpm which is almost unaffected, the curve in gray shows the GSR signal, the curve in yellow the filtered phasic GSR and the blue curve the detected GSR peaks. We see an initial rise in the GSR signal at the beginning which could either be caused by the demand of the driving task in general or the stress of the particular experiment section. The concentration of peaks at the first half of the curve might indicate the former for the parking exit but the latter for the curve section. The participant seems to relax after that until the dynamic intersection where we see an increase in GSR level and number of GSR peaks.

---

Figure C.9 shows the time sequence of a normal driver viewpoint condition that turned out less smooth and took over 9 minutes. The participant rotates his head a lot to gain awareness of the surrounding situation as indicated by the head direction graph, reaching a peak of more than 120 degrees during parallel parking where the participant attempted to look backwards through the rear window. In the first half of the curve section we see that the car heading graph is more complex and includes a moment of straight driving towards west. In the dynamic intersection, we observe a crash with cross traffic from right in repetition #2 at around  $x=1500$  indicated by the jump in car heading to the left. We see yet another crash in repetition #3 in which the heading changed less abruptly as the car was dragged along by the cross traffic. Finally, we see another crash at the beginning of repetition #4, this time caused by a stop location which was too much forward into the intersection, followed by a backing maneuver to gain enough safe distance from cross traffic. HR is again unaffected even from these crashes. GSR levels raise belated after the first crash, immediately before the second when the participant noticed the dangerous situation and finally once again in repetition #4 where he saw another dangerous situation which resulted not in an accident but a near-miss event.

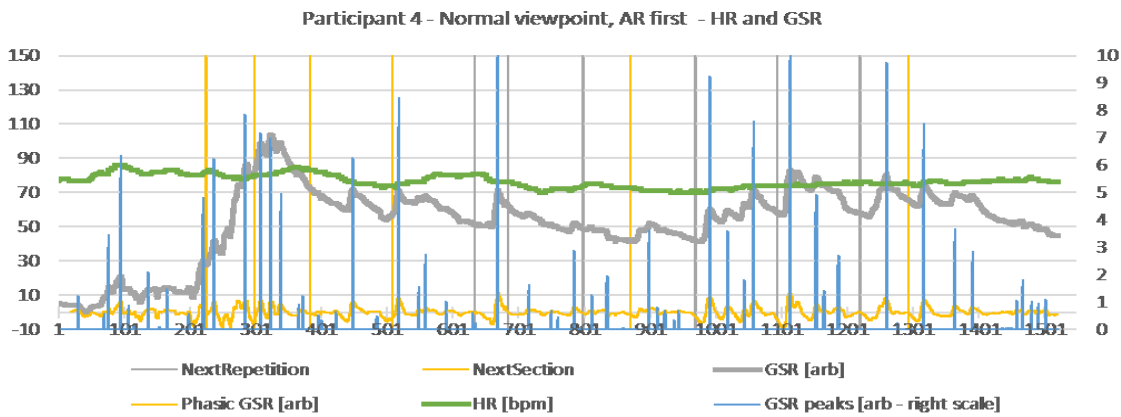
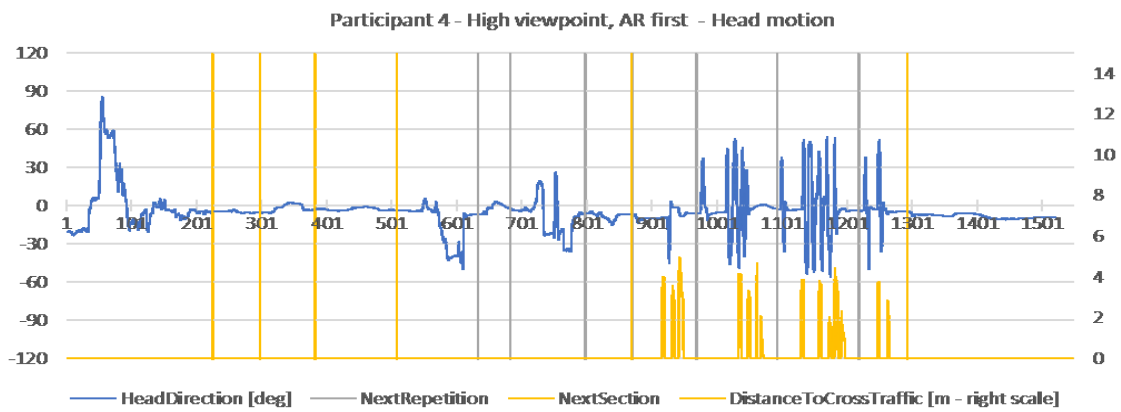
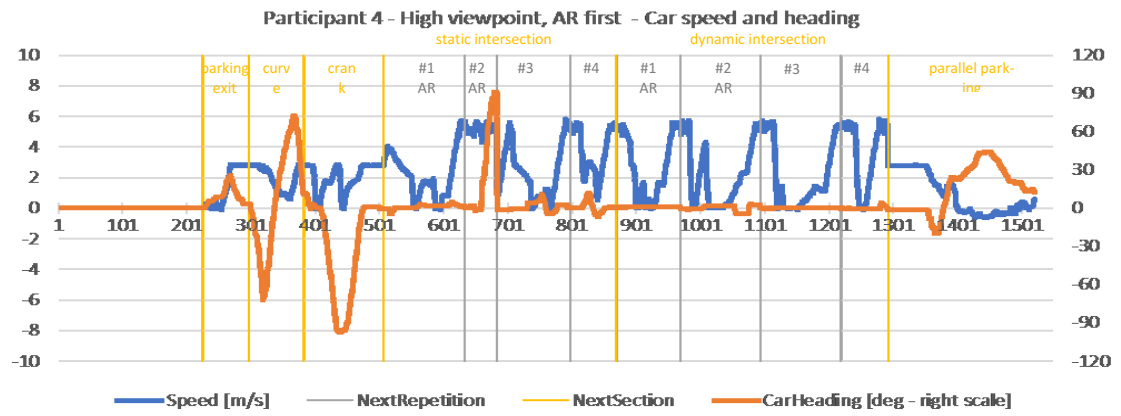


Figure C.8: Time sequence of high viewpoint condition for participant 4.



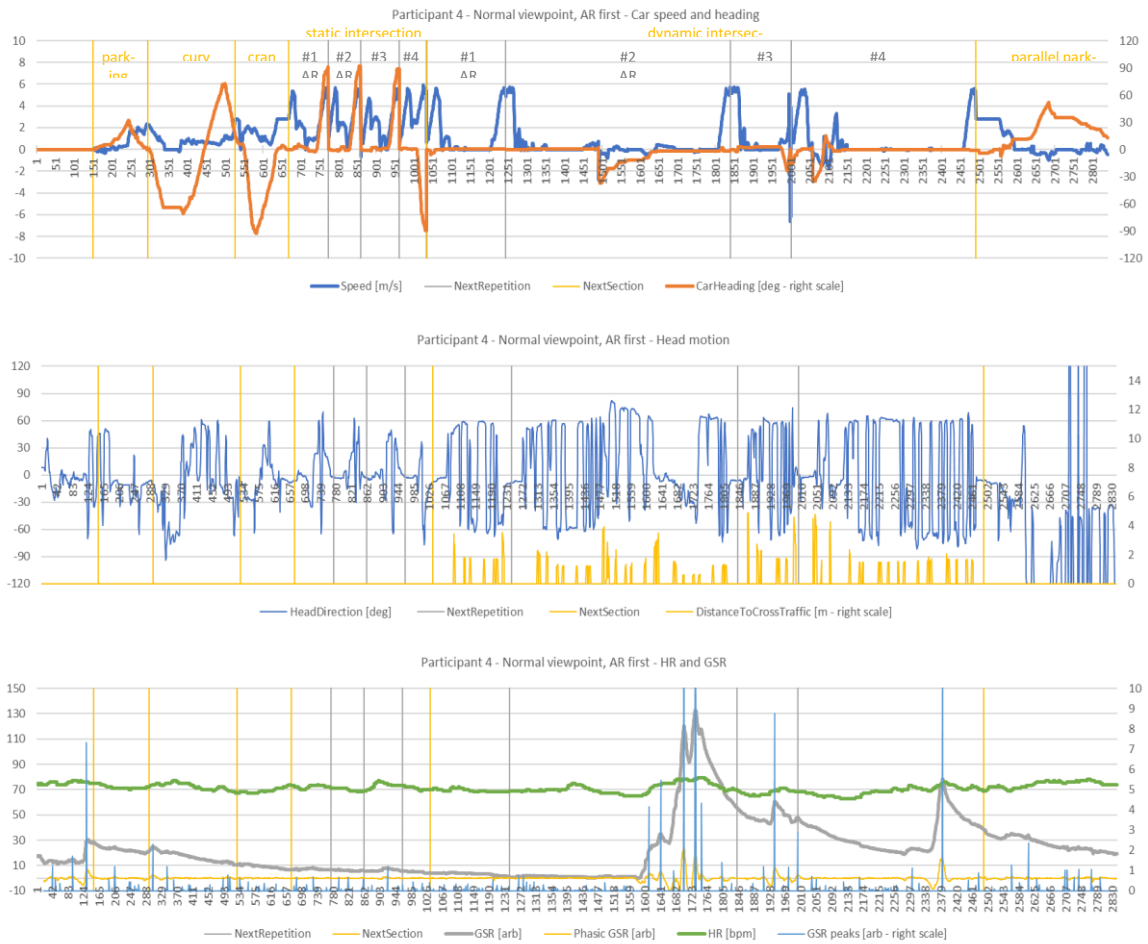


Figure C.9: Time sequence of normal viewpoint condition for participant 4.

---

## C.6 GSR

GSR data from 3 subjects were discarded due to noise, signal clipping and unclear data, leaving us with data from 3 subjects. While subject 4 indicated no motions sickness at all, subject 5 experienced motions sickness in the later conditions and subject 6 in all conditions. This might have influenced the GSR signal.

For subject 4, the number of GSR peaks was lower for high viewpoint than for normal viewpoint in all sections, and for low viewpoint in all sections except crank. For subject 5, the number of GSR peaks was similar between the three experiment conditions in all experiment sections with the exception of normal viewpoint in the dynamic intersection, possibly related with the 2 near misses vs 1 near miss in the other conditions. For subject 6, the number of GSR peaks was lower for high viewpoint compared with normal viewpoint for parking exit, dynamic intersection and parking but not for curve, crank and static intersection. Low viewpoint was lower than normal only for static intersection and parking but higher in all other sections.

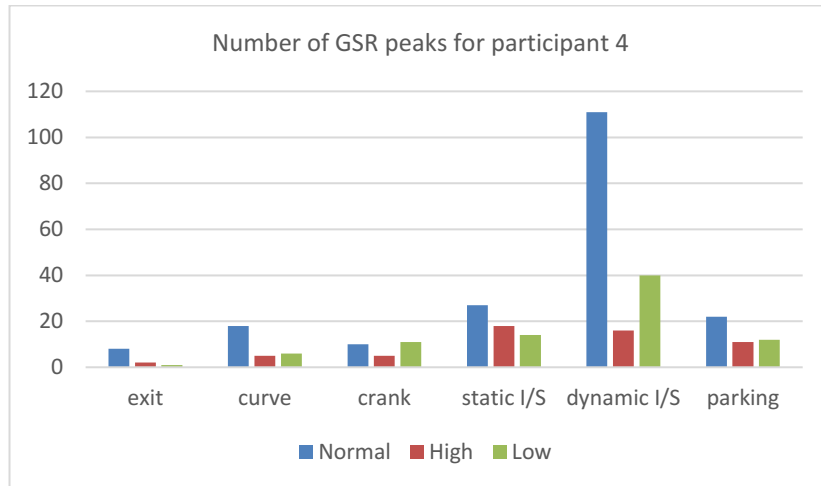


Figure C.10: Number of GSR peaks for participant 4.

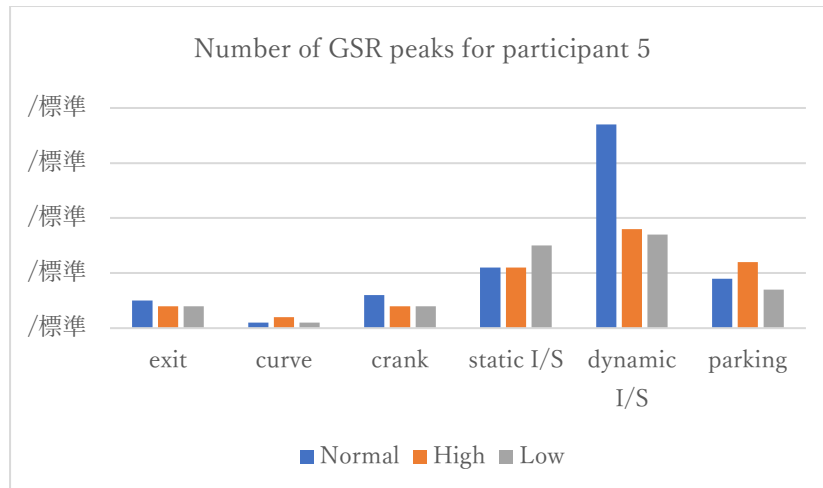


Figure C.11: Number of GSR peaks for participant 5.

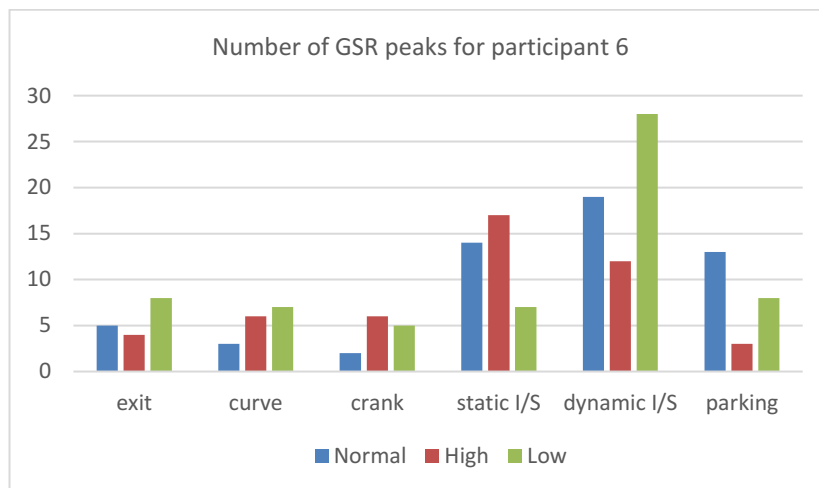


Figure C.12: Number of GSR peaks for participant 6.

### C.7 HR and HRV

Data of 3 subjects were discarded due to malfunction of sensor strap, leaving us with data from 3 subjects. We could not identify any clear trends from HR and HRV data.

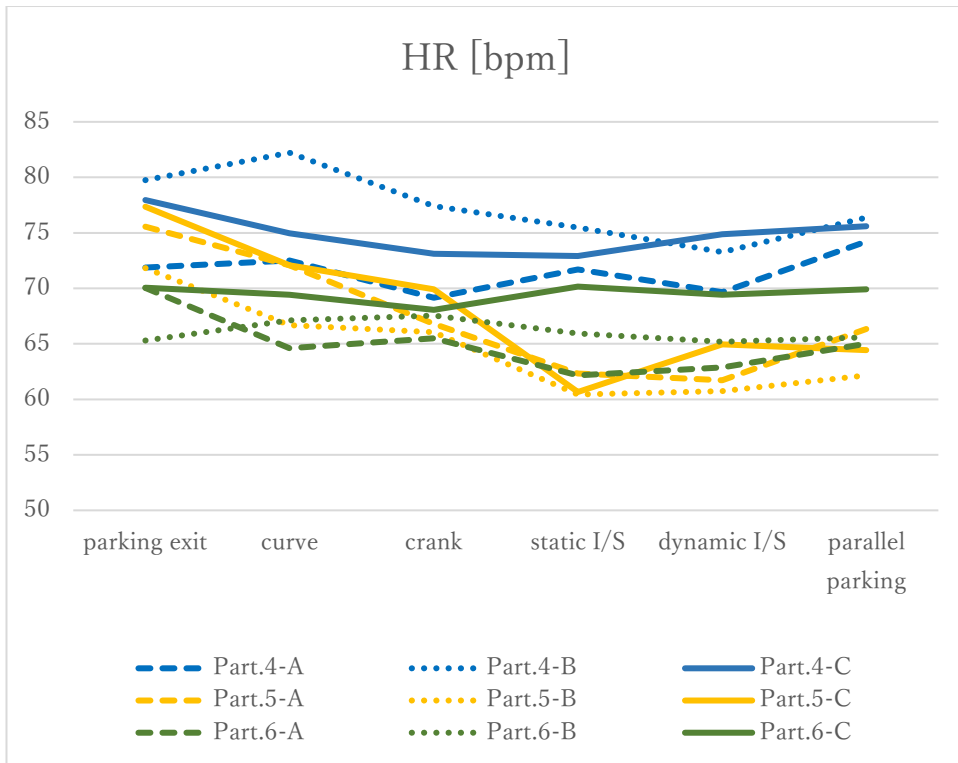


Figure C.13: HR.

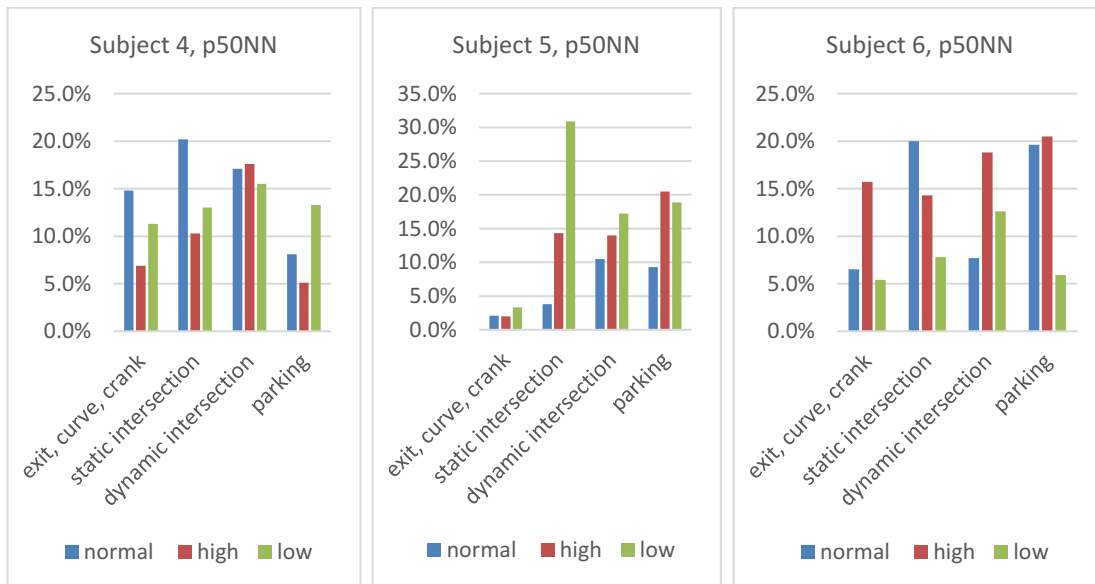


Figure C.14: HRV data (p50NN).

## C.8 Detailed analysis results

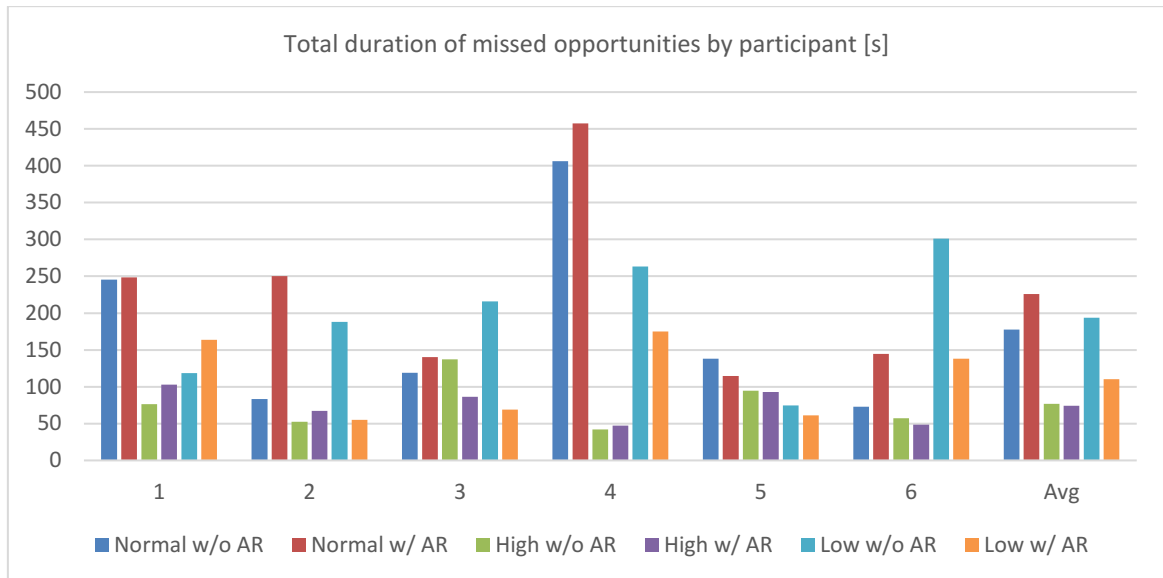


Figure C.15: Total duration of missed opportunities by participant.

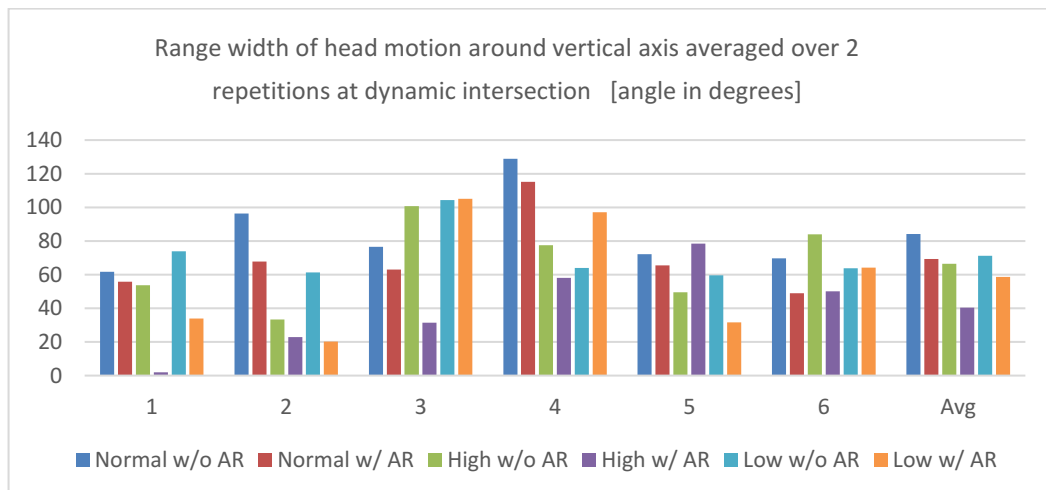


Figure C.16: Range of head motion around vertical axis at dynamic intersection.

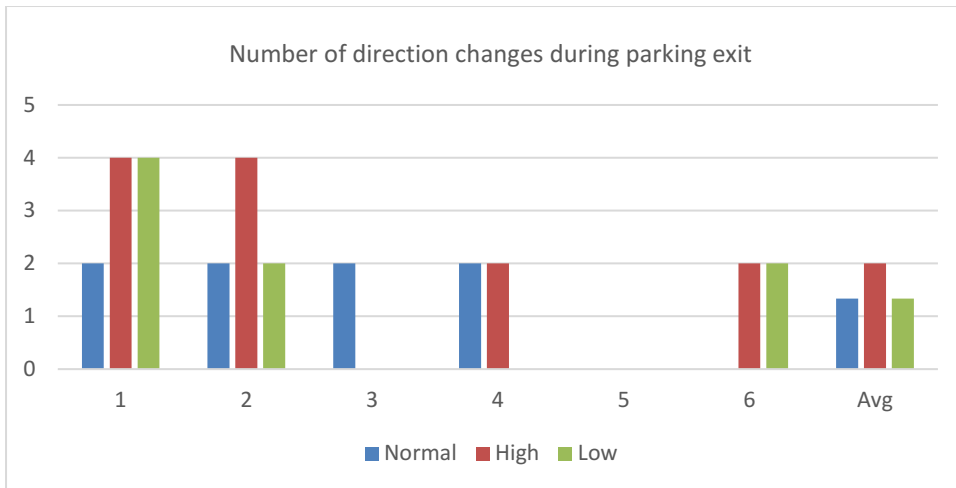


Figure C.17: Number of direction changes during parking exit.

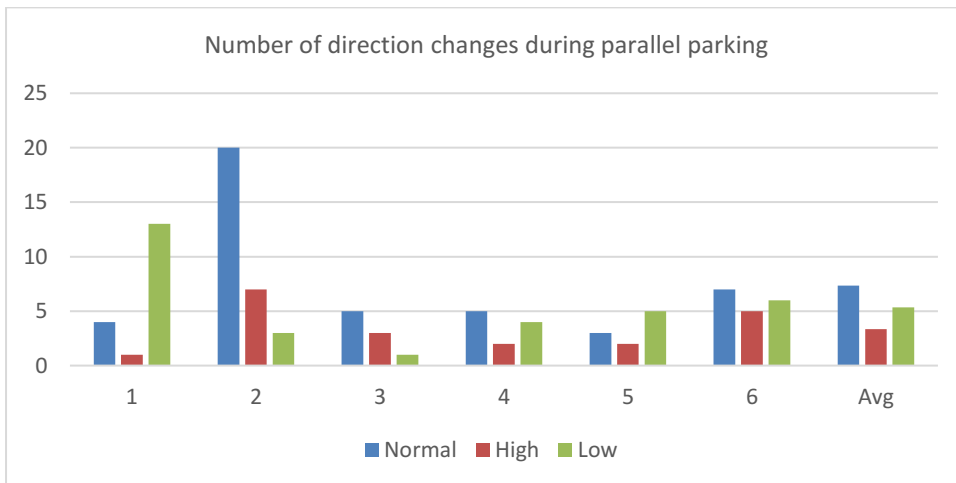


Figure C.18: Number of direction changes during parallel parking.

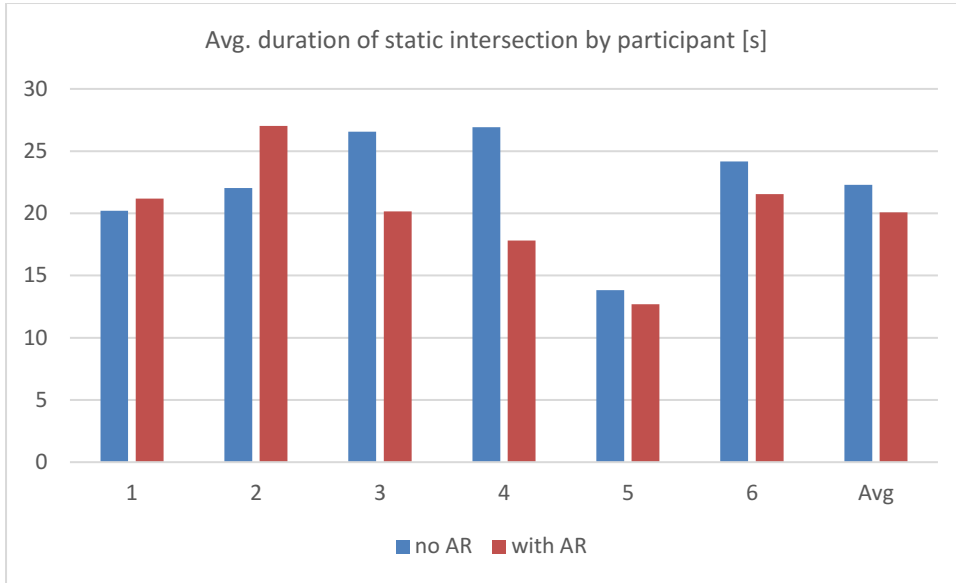


Figure C.19: Duration of static intersection by AR.

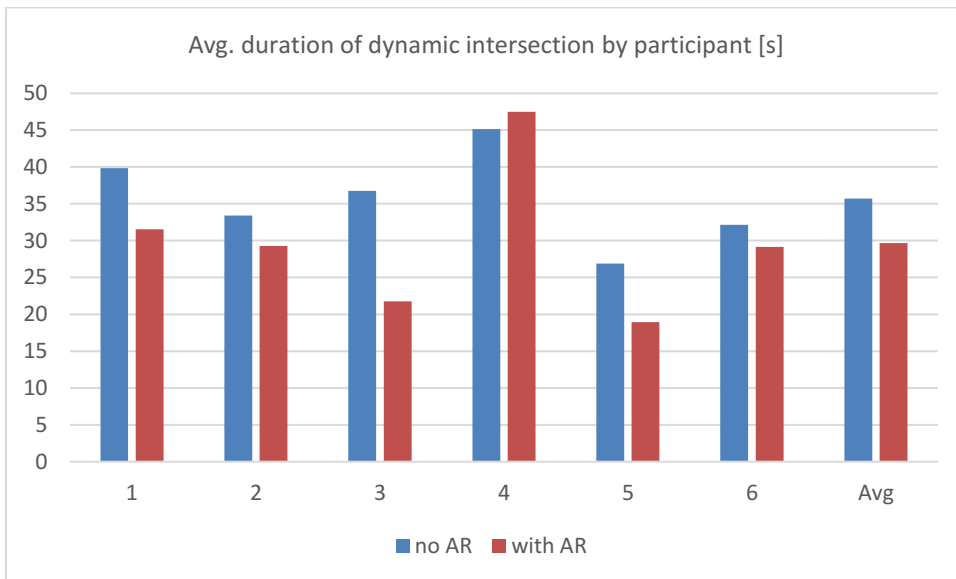


Figure C.20: Duration of dynamic intersection by AR.

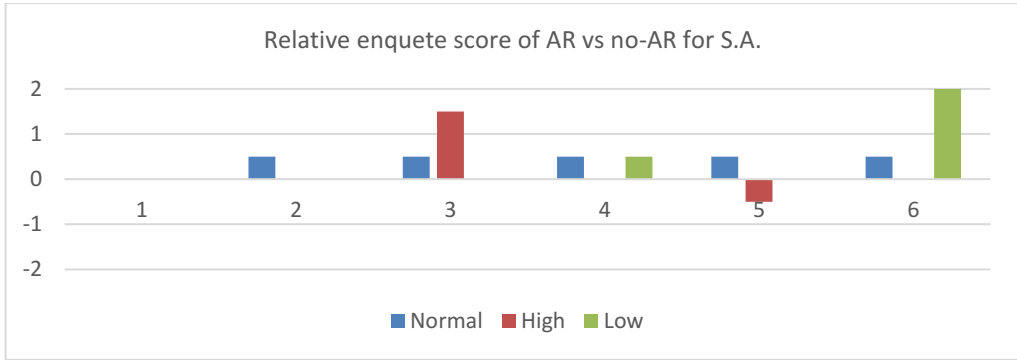


Figure C.21: Relative enqueue score comparing AR against no-AR by participant.

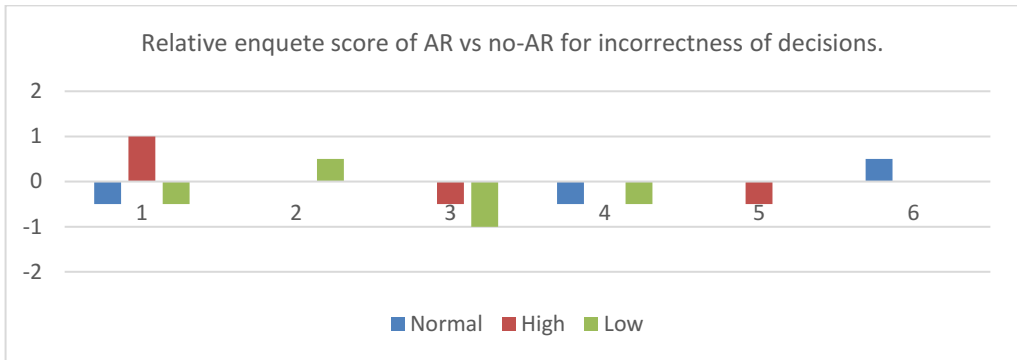


Figure C.22: Relative enqueue score comparing AR against no-AR by participant.

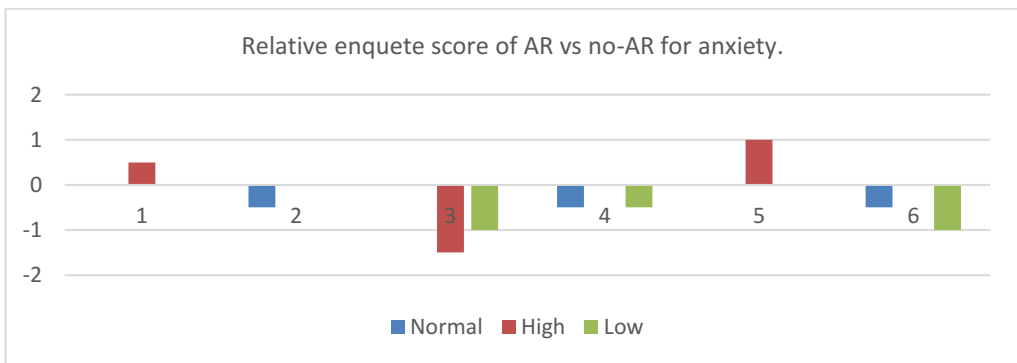


Figure C.23: Relative enqueue score comparing AR against no-AR by participant.



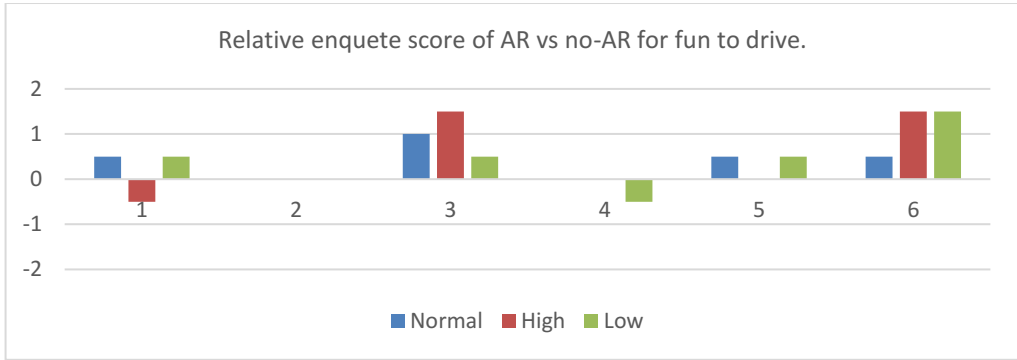


Figure C.24: Relative enquete score comparing AR against no-AR by participant.

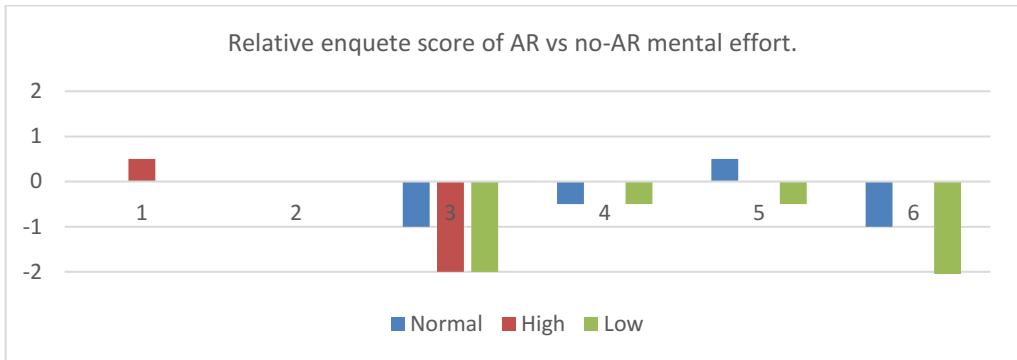


Figure C.25: Relative enquete score comparing AR against no-AR by participant.

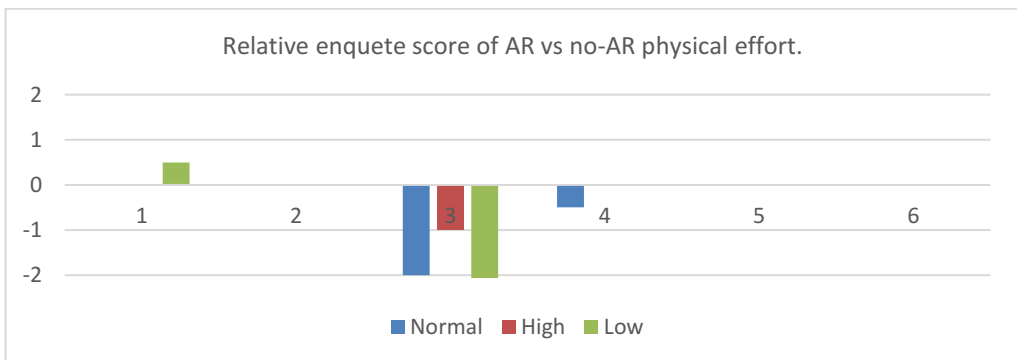


Figure C.26: Relative enquete score comparing AR against no-AR by participant.

## C.9 Enquete form

アンケート [1/3] 実験開始日時：

氏名：  
 年齢：  
 性別：  
 運転歴： 年（実際に運転している期間、ペーパードライバー期間を除く）  
 運転頻度（過去1年間の平均）： 日/週 回/週 時間/週

下記の質問に5段階で回答をお願いします。

とても当てはまる                      どちらでもない                      全く当てはまらない  
 |-----|-----|-----|-----|-----|

本アンケートにおける言葉の意味。

「状況把握する」とは、運転している車両の周囲の道路形状、交通ルール、および、車両・自転車・歩行者などの存在・位置・動きを把握することを言います。

「正しく運転する」とは、車道からはみ出ず、道路上・道路外の障害物や車両や歩行者などに接触せずに、進入禁止・一旦停止・優先順位などの法規を守って運転することを言います。

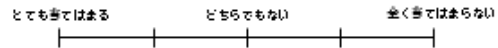
### 実験前

#### 各シーンにおける自分の運転能力について

	状況把握は 得意だ	正しく運転 するのは苦手だ	運転に不安を 感じる
縦列駐車発進	-----	-----	-----
カーブ	-----	-----	-----
クランク	-----	-----	-----
交差点・経路判断	-----	-----	-----
交差点・速度判断	-----	-----	-----
縦列駐車	-----	-----	-----
	運転が楽しい	心理的負荷が 高い	体力的負荷が 高い
縦列駐車発進	-----	-----	-----
カーブ	-----	-----	-----
クランク	-----	-----	-----
交差点・経路判断	-----	-----	-----
交差点・速度判断	-----	-----	-----
縦列駐車	-----	-----	-----

アンケート [2/3]

氏名:



### 実験条件 1

( )

**AR**→**AR**

**AR**→**AR**

もし酔いを感じた場合は

	こ ま ま	こ ま ま	こ ま ま
縦列駐車発進			
カーブ			
クランク			
交差点・経路判断			
交差点・経路判断			
縦列駐車			

	状況把握 できていた	正しく 運転できなかった	運転に不安を 感じた
縦列駐車発進			
カーブ			
クランク			
交差点・経路・ <b>AR</b>			
交差点・経路・ <b>AR</b>			
交差点・速度・ <b>AR</b>			
交差点・速度・ <b>AR</b>			
縦列駐車			
縦列駐車発進			
カーブ			
クランク			
交差点・経路・ <b>AR</b>			
交差点・経路・ <b>AR</b>			
交差点・速度・ <b>AR</b>			
交差点・速度・ <b>AR</b>			
縦列駐車			
縦列駐車発進			
カーブ			
クランク			
交差点・経路・ <b>AR</b>			
交差点・経路・ <b>AR</b>			
交差点・速度・ <b>AR</b>			
交差点・速度・ <b>AR</b>			
縦列駐車			

### 実験条件 2

( )

**AR**→**AR**

**AR**→**AR**

もし酔いを感じた場合は

	こ ま ま	こ ま ま	こ ま ま
縦列駐車発進			
カーブ			
クランク			
交差点・経路判断			
交差点・経路判断			
縦列駐車			

	状況把握 できていた	正しく 運転できなかった	運転に不安を 感じた
縦列駐車発進			
カーブ			
クランク			
交差点・経路・ <b>AR</b>			
交差点・経路・ <b>AR</b>			
交差点・速度・ <b>AR</b>			
交差点・速度・ <b>AR</b>			
縦列駐車			
縦列駐車発進			
カーブ			
クランク			
交差点・経路・ <b>AR</b>			
交差点・経路・ <b>AR</b>			
交差点・速度・ <b>AR</b>			
交差点・速度・ <b>AR</b>			
縦列駐車			
縦列駐車発進			
カーブ			
クランク			
交差点・経路・ <b>AR</b>			
交差点・経路・ <b>AR</b>			
交差点・速度・ <b>AR</b>			
交差点・速度・ <b>AR</b>			
縦列駐車			

アンケート [3/3]  
氏名:

**実験条件 3**

( )

**AR→AR**

**AR→AR**

もし酔いを感じた場合は

	こ 少な	こ 多
縦列駐車発進		
カーブ		
クランク		
交差点・経路判断		
交差点・経路判断		
縦列駐車		

	とても当てはまる	どちらでもない	全く当てはまらない
	----- ----- -----		
	状況把握 できていた	正しく 運転できなかった	運転に不安を 感じた
縦列駐車発進			
カーブ			
クランク			
交差点・経路・ <del>AR</del>			
交差点・経路・ <b>AR</b>			
交差点・速度・ <del>AR</del>			
交差点・速度・ <b>AR</b>			
縦列駐車			
	運転が楽しかった	心理的負荷が 高かった	体力的負荷が 高かった
縦列駐車発進			
カーブ			
クランク			
交差点・経路・ <del>AR</del>			
交差点・経路・ <b>AR</b>			
交差点・速度・ <del>AR</del>			
交差点・速度・ <b>AR</b>			
縦列駐車			

**実験条件 4**

( )

**AR→AR**

**AR→AR**

もし酔いを感じた場合は

	こ 少な	こ 多
縦列駐車発進		
カーブ		
クランク		
交差点・経路判断		
交差点・経路判断		
縦列駐車		

	とても当てはまる	どちらでもない	全く当てはまらない
	----- ----- -----		
	状況把握 できていた	正しく 運転できなかった	運転に不安を 感じた
縦列駐車発進			
カーブ			
クランク			
交差点・経路・ <del>AR</del>			
交差点・経路・ <b>AR</b>			
交差点・速度・ <del>AR</del>			
交差点・速度・ <b>AR</b>			
縦列駐車			
	運転が楽しかった	心理的負荷が 高かった	体力的負荷が 高かった
縦列駐車発進			
カーブ			
クランク			
交差点・経路・ <del>AR</del>			
交差点・経路・ <b>AR</b>			
交差点・速度・ <del>AR</del>			
交差点・速度・ <b>AR</b>			
縦列駐車			

ご協力ありがとうございました。

## C.10 Verbal comments

Subjects were asked to verbally comment their subjective impressions during the experiment.

Table C.7 : Verbal comments

Condition	Phase	Subject					
		1	2	3	4	5	6
Driver	Parking exit			やっぱり（障害物までの）感覚がつかみづらい	後ろの距離判断が難しい		
	Curve, crank	この視点難しい	車幅の感覚がよくわからない	当たっているかわからない	Detailよりみやすい 左側の距離感がわからない。S字よりはクランクが行きやすい…でも当たったかも。		
	Static intersection			看板は上からよりも見やすいですね	Driver、Detailはあまり変わらない		
	Dynamic intersection	えっ..えっ、あー		上からの方がわかりやすいですがこっちの方が慣れている。ずっと右見ていればいいんだ (Routine)	自車の先端の長さがわからない		
	Parallel parking	これ難しい	ええっ、やべっいや、難しいな、ちょっとつかっています		ルームミラー・左ドアミラーが使えないのはとても難しく		
Overview	Parking exit						Detailより発進しやすい
	Curve, crank	めっちゃ凄いや、ゲーム感覚		上から見えるのは使いやすいい…ぶつかっていますけど	難しい、当たる…OK… 回転半径をあまり把握できていない これがベスト		ハンドルの切り始めタイミングが掴みやすくてよい。ギリギリを責められるのでゲームみたいに楽しい。
	Static intersection	この視点はいいですね、わかりやすい、サクサク行ける感じが	わ、間違えちゃいました	上からだ左右の標識は見づらい	これがワースト（標識が見づらい）		
	Dynamic intersection	これすげえ便利だなうわ、これはすげえいい、めっちゃ見やすい		これは上から見えるのはとてもいいです、もっと上からでもいいかも	視点が高いのは楽		急に視点の高さが変わると不思議な感じ
	Parallel parking	おお、この視点はすごい運転しやすい		これも見やすい、わかりやすい			
Detail	Parking exit						
	Curve, crank			Driverより見やすいのかな、（ボールへの）距離がつかみやすい気がします	あっ…		クランク難しい。左端の車幅感覚がわからない
	Static intersection	ドライバーの視点と大差ない		看板が見づらい、のぞき込まないといけない	特に難しくない		
	Dynamic intersection				難易度が一気に上がった		
	Parallel parking	これ難しい		（リアビューは）いいですね…ああ	後方視点は難しい（左右反転していないから？）		
AR	Static intersection			ARがあるとそっちを見ないで済むので楽ですね (Overview、Driver) 注意する箇所が2か所に減るのはありがたい (Detail)			ARがあった方がわかりやすい
	Dynamic intersection	便利 そっか、赤いのはぶつかってもいいのか	そっか、AR表示はぶつかってもいいのか	これは何も考えなくていいので便利 (Overview) ARがあってもちゃんと見たい、慣れてくるとよいかも (Driver) 信じられるようになればきつとすごい便利 (Detail)	ARありなしのやりやすさがあまり変わらない…でもあった方が…やっぱりあった方が微妙な隙間などでありがたい (Detail) ARあった方が渡りやすい。右からと左からのARは独立しているの？ 車が多いとARがないと難しい。普段ならこんな道は渡れない。 (Driver)		
Other comments	HMD上げえ、後ろも見える コントローラ難しい	標識が見づらい（解像度） 自車の加速力がわからないので（条件5は）難しい インチャングが難しい（条件5） Overviewの方が視線移動がすくなくてHMD酔いしづらい		コントローラは使いづらい 看板覚えていないと難しい			Overviewで下を向くと酔う