CHALMERS

The Design of a Multi-touch Gestural System for Semi-autonomous Driving

Master of Science Thesis

JOEL HAMMAR
ANDREAS KARLSSON

Department of Applied Information Technology
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden, 2015

Report No. 2015:123

Report Number: 2015:123

The Design of a Multi-touch Gestural System for Semi-autonomous Driving

JOEL HAMMAR
ANDREAS KARLSSON

Department of Applied Information Technology
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden, 2015

The Design of a Multi-touch Gestural
System for Semi-autonomous Driving
JOEL HAMMAR
ANDREAS KARLSSON

Master thesis at Chalmers University of Technology
In cooperation with Semcon
Report No. 2015:123

Department of Applied Information Technology
Chalmers University of Technology
SE-412 96 Gothenburg, Sweden
Telephone + 46 31-772 1000

Cover:
The prototyped gesture system, placed inside a vehicle.

Gothenburg, Sweden, 2015

Abstract

With more and more advanced technology being developed, the need of human interaction is no longer as prominent as before. Especially in the automotive industry where autonomous vehicles come closer to being introduced to the consumer market. However, by introducing automation the role of the driver is changing from being the operator of the vehicle to becoming a supervisor of the automation state (Damböck et al. 2012). That can lead to potential problems where the driver becomes less aware of the current situation and the reaction time in case of system failure is increased (Endsley and Kiris 1995; Neimann, Peterman, and Manzey 2011). In order to address this problem we have designed a multi-touch gestural system for semi-autonomous driving, which enables the driver to influence the driving by giving instructions, such as change lane or take next exit/turn to the vehicle while it is in autonomous mode.

The process towards designing the system included several iterations. In the first and second iteration the focus was on understanding the domain of gestures and which gestures that are most natural for giving the vehicle instructions. Then in the third iteration the focus progressed towards designing multimodal feedback and evaluating the complete system in a high-fidelity simulator. Lastly, a few changes were made in the design of the system according to the data gathered from the evaluation.

Conclusions that can be drawn from this study is that the users seem to prefer gestures which aims to resemble the movement of the vehicle. Furthermore, the results indicate the single finger gestures are prefered rather than multi-touch and that the gestures should be designed to be as simple as possible in terms of gestural motion. Additionally, the feedback should be instantaneous and multimodal (visual and auditory), as multimodality seem to convey a richer message to the user than if the modalities were to be used separately.

Keywords: gestures, multimodal feedback, semi-autonomous driving, automated vehicle, interaction design, human-machine interaction, swipe.

Acknowledgements

This study was carried out as a master thesis at the Interaction Design & Technologies programme at Chalmers University of Technology in cooperation with Semcon, Volvo Cars and Viktoria Swedish ICT. We would like to thank Semcon for providing the opportunity to carry out this study and especially our external supervisor Jan Nilsson for great support throughout the study. Additionally, we would like to thank Claes Edgren at Volvo Cars for his feedback and for providing us with test vehicles.

Foremost, we would like to express our gratitude to our supervisor Jonas Landgren for his commitment and immense support throughout the study. We could not have had a better supervisor. His guidance and expertise has been of great help for us.

Lastly, we want to thank everyone who has taken part in our user studies or in other ways been involved in the study. We highly appreciate all the time, effort and input that they have contributed to the study.

Gothenburg, June 2015
Joel Hammar and Andreas Karlsson

Table of Contents

1. Introduction

As technology advances more and more products become autonomous. Vacuum cleaners are able to vacuum an entire room by themselves, home climate control can automatically adapt to families habits, and blinds can auto fold whenever the sun is too bright. This may seem very convenient. Yet, making everything autonomous omit people from being able to influence the behaviour. Similarly, autonomous vehicles exclude the driver's ability to influence the driving. Therefore a need to make use of both the convenience from the autonomous world and still keep the possbility for humans to interact have emerged. In this borderland between humans and machines, semi-automated systems challenge us to design useful, usable and intuitive Human Machine Interfaces (HMI).

An area in which an increasing effort on autonomation can be observed is the vehicle industry, where even companies outside the industry are entering the competition of introducing the first autonomous vehicle to the consumer market. For instance, Google have been testing fully functional autonomous cars on enclosed test tracks and will start testing them on streets in California in the beginning of 2015 (Abrahamsson, 2013; Google 2014). Further, Apple recently began researching this area as well (Taylor and Oreskovic 2015). Additionally, in 2017 Volvo will have up to 100 cars driving autonomously on the highway roads around Gothenburg (Volvo Cars 2013).

It is evident that the current development in the vehicle industry is moving towards introducing fully automated vehicles. However, these vehicles omit the driver from being able to influence the driving. Hence, a natural next step would be to introduce semi-autonomous vehicles in which the benefits of the autonomous world and the benefits of the human's abilities can be combined. A currently ongoing research project investigating the field of semiautonomous vehicles is the AIMMIT project, which is run by Volvo Cars in cooperation with Semcon, Viktoria Swedish ICT and Chalmers University of Technology. The aim of the project is to *"explore the potential benefits of different multimodal HMI concepts for current and future vehicle functionality in terms of safety, customer satisfaction and competitiveness"* (Chalmers 2015). This study is a part of the AIMMIT project and will explore the possibilities of novel interaction modalities for semi-autonomous driving.

1.1. Research Problem

As the level of automation is increasing, the role of the driver is changing from being the operator of the vehicle to becoming a supervisor of the automation state (Damböck et al. 2012). Thus, the amount of involvement required from the driver is decreased, which can lead to potential *out-of-the-loop* performance problems where the driver becomes less aware of the current situation and lose manual driving skills (Endsley and Kiris 1995; Niemann, Peterman, and Manzey 2011). This is one of the main problems with introducing automation in vehicles (Gold et al. 2013).

To solve this problem the driver should be kept *in-the-loop* and aware of the current situation (Endsley and Kiris 1995; Niemann, Peterman and Manzey 2011) regarding the environment around the vehicle such as the state of the road, the weather conditions, and surrounding vehicles etc. Endsley (1988 p. 792) defined situational awareness to be *"the perception of elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future"*. If the driver would fall out-of-the-loop it can impair his or her ability to manually take over the control of the vehicle in case of automation failure (Endsley and Kiris 1995; Niemann, Peterman, and Manzey 2011; Gold et al. 2013), potentially leading to dangerous situations.

1.2. Research Question

One of several possible alternatives for keeping the driver in-the-loop is to implement a system which enables the driver to influence the task of driving by giving instructions to the vehicle such as change lane, take the next exit/turn, increase speed or overtake. The interaction between the system and the driver could be designed in many ways of which one of the least obtrusive implementations could be a touch-based gesture system. To design and implement such a system for semi-autonomous driving is the main objective for this study. With that, novel interaction modalities are introduced and the traditional interactions for the task of driving are omitted. Thus, challenges in the design work can be particularly difficult since it may require solving problems with a high amount of uncertainty and complexity. Therefore, the study will be explorative and seek to investigate which properties a gestural interface for semi-autonomous driving should have. The research question is formulated as follows:

> *Which properties should a multi-touch gestural system for semi-autonomous driving have?*

1.3. Aim

The study seeks to investigate which opportunities and potential issues a multi-touch gesture based system can bring to semi-autonomous driving. Furthermore, it aims to explore which input gestures that are most natural for semi-autonomous driving. Lastly, it aims to provide a design solution for how the feedback can be presented to the driver.

1.4. Delimitations

The study does not include research about which instructions the driver should be able to give to the vehicle via the input device. These have been predefined by the stakeholders before the study. Moreover, the system will not be designed to handle a take over from the driver or communication with the GPS. Lastly, the study will not investigate how the system can be designed for disabled people nor will the solution be adapted to anatomical differences among users.

1.5. Scope

The scope of the study lies within semi-autonomous driving (see Fig. 1) in the middle of the spectrum from manual driving to fully automated driving (Scheiben et al. 2011). In semi-autonomous vehicles the task of driving is partly automated, which means that the vehicle is able to drive all by itself but the driver may be required to take over the control if the situation demands that. Furthermore, as the driver is able to influence the task of driving while the vehicle is in autonomous mode, it can not be considered as fully automated driving. Neither is it manual driving or driver assisted, which defines the level of automation to semi-autonomous. Thus, manual driving, assisted driving, highly autonomous driving and “driverless cars” (i.e. fully automated) lie outside the scope of this study.

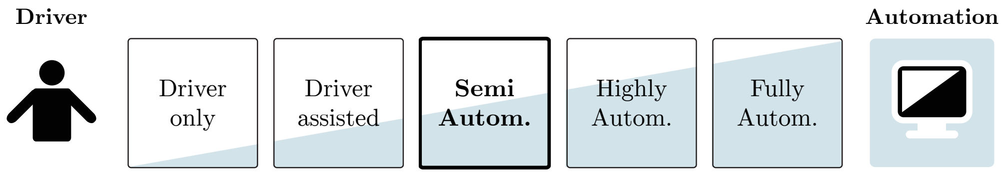


Figure 1. The spectrum of different levels of automation where our study lies within the scope of semi-automated driving. The figure is adapted from Scheiben et al. (2011).

Within semi-autonomous driving the design space within the vehicle will originate from the driver's position in the vehicle (i.e. passengers will not be considered). Within this design space, there are four different dimensions which we identified early in the study. These four dimensions are gestures, feedback, placement and form of the input device. The first two dimensions refers to which touch gestures that are most natural for giving instructions to the vehicle and how the feedback should be presented to the driver. The later two regards the form of the input device itself and where it should be placed in the vehicle. However, these two dimensions will not be in focus of the study though they will be considered in the design.

1.6. Stakeholders

This study was conducted in cooperation with Semcon AB, Volvo Cars AB, Viktoria Swedish ICT and Chalmers University of Technology. The research was carried out at the Semcon office in Gothenburg with additional user tests and evaluations in the field. Semcon was the main stakeholder and provided the main objectives for the study. Semcon is an international technology company in engineering services and product information. Their services include product development, product information, quality control, training and methodology development within fields such as automotive, life science, telecommunications, energy etc. The head office in Gothenburg has a strong focus towards the automotive industry which is one of their specialist areas in the field of user experience and interaction design (Semcon 2015).

2. Related Work

The field of automation have been a subject for research for a long period of time, as the first autopilots were introduced back in the 1920s (Billings 1996), almost 100 years ago. In the aviation industry, similar problems with pilots falling out-of-the-loop have been observed (Billings 1997). According to Billings (1997), pilots should be involved in maneuvering the aeroplane even when the autopilot is turned on in order to stay in-the-loop. Similar results have been presented by Niemann, Peterman and Manzey (2011), but within the automotive domain. They designed a solution which aimed to prevent out-of-the-loop problems for a vehicle driver. Their study indicated that by letting the driver delegate tasks to the vehicle while it is in autonomous mode, the driver's situational awareness can be increased and the reaction time in case of system failure can be shortened. In other words, if the driver is involved in the task of driving he or she is more likely to stay in-the-loop. Further along out-of-the-loop problems, Endsley and Kiris (1995) studied decision making for a navigation task in different levels of automation. Their results indicated that as the level of automation increases, the operator's situational awareness decreases and the reaction time in case of system failure becomes longer. That seem to be due to the shift from active to passive information processing when automation is introduced. Thus, they concluded that out-of-the-loop performance problems could be minimized by keeping the operator active in automated conditions. Further along this subject, Damböck et al. (2012) studied how status communication in highly autonomous vehicles can be used to increase the driver's understanding of the vehicle's actions. Their study pointed towards that if the driver is given visual feedback about what the vehicle is planning to do next, the cooperation between the driver and the vehicle could be enhanced.

Moving from the domain of autonomous driving to the the area of gestural interfaces, we can see that lots of research have been made both within and outside the automotive domain. However, none of it include gestural interfaces in autonomous vehicles. Hence, we have reviewed research regarding gestural interaction more broadly related to our work. Starting outside the automotive domain, Hinrichs and Carpendale (2011) studied how people use multi-touch gestures on a large interactive table. The main finding from the study was that a variety of gestures may be natural for any given intended action and which gesture that is used is influenced by the interaction context and the social context. Furthermore, different people tend to use different variations of a gesture for the same action. Continuing on gestural interfaces, Fang and Ainsworth (2012) investigated how multimedia functions in vehicles could be controlled using touch gestures performed on the steering wheel. In their study they created a prototype of a steering wheel with seamlessly integrated touch pads. That created some issues for the users because they had troubles understanding where the input device was placed and how it could be used. This issue was not present in the study by Rahman, Saboune and Saddik (2011) who also studied how multimedia functions can be controlled, but by using free-form gestures instead of touch gestures. Their result indicated that by using auditory and haptic feedback (vibrating bracelet) it can make drivers feel less distracted and decrease their visual load. Further along the line of controlling infotainment using gestural interaction, Andersson and Wikander (2013) studied how touch-based gestural interaction can used to minimize visual distraction for the driver. Their work resulted in a set of guidelines for gestural interfaces. They suggest that one should conform to conventions if possible, and if not it is important to test the gestures with users. In addition, they argue that patterns are important for a set of gestures since it will increase the coherency and facilitate remembering the gestures. Continuing on multi-touch interaction, Pfleging et al. (2011) investigated possible advantages of using multi-touch combined with other modalities. They concluded that by taking advantage of multimodal interaction, potential drawbacks for each single modality can be decreased. Lastly on the subject of multi-touch interaction, Butz, Kramer and Rümmelin (2013) conducted a study on performance difference between multi-touch gestural interfaces and button interfaces. Their result indicated that multi-touch gestural interfaces outperforms button interfaces since it is equal in efficiency but also lowers the workload for the user. However, the multi-touch gestures might require a learning phase, but after that the usability is increased compared to button interfaces.

In addition to the research on gestural interfaces, we have reviewed studies on methodology for designing gestures and gesture hinting. The first one by Morris, Wobbrock and Wilson (2009) aimed to investigate if gestures designed by HMI experts or users would be prefered by end users. Their result indicated that users prefer to use one hand over two hands, that they do not care so much about how many fingers they use. However, in a later study in 2010, Morris, Wobbrock and Wilson concluded that users prefer using only a single finger when performing a gesture which contrasts to their previous findings. Furthermore, the users generally found analog gestures more intuitive than abstract and prefered gestures that required less cognitive effort and were physically easy to perform. Lastly, the authors concluded that gestures developed together with users are more easily understood and prefered over gestures that are designed only by HMI researchers. Similar results have also been presented by Nielsen et al. (2003) who conducted a study comparing a technological approach with a user centred approach for designing gestural interfaces. Their result indicated that gestures designed by users are better compared to if the gestures are designed based on the available technology. Lastly, we reviewed a study on the subject of gesture hinting by Lundgren and Hjulström (2011). They state that it can be problematic for users to discover possible interactions in mobile gesture interfaces. To solve that, they propose a set of symbols which can be used as dynamic gesture hints on the screen describing possible interactions.

After reviewing this research we can see that our work continues along the line of out-of-the-loop problems studied by Endesly and Kiris (1995), Billings (1997), and Niemann, Peterman and Manzey (2011). The system we will design will allow the driver to influence the driving while the vehicle is in autonomous mode, which could reduce out-of-the-loop problems. Furthermore, by communicating appropriate feedback about what the vehicle will do next to the driver, it could make the driver more involved in the task of driving and aware of the current situation. Additionally, the feedback could improve the cooperation between the driver and the vehicle, similar to the result of the study by Damböck et al. (2012). However, we will not investigate these out-of-the-loop problems explicitly as it is not the focus of our study. Instead, we will focus on designing an in-vehicle gestural interface, which has been previously studied by Andersson and Wikander (2013), Rahman, Saboune and Saddik (2011), Fang and Ainsworth (2012) and Pfleging et al. (2011). However, the scope of our study is different since the system we design will be used for semi-autonomous driving and not manual driving. Thus, factors such as driver's distraction will not be an issue in the same manner as in the previous studies. Furthermore, we will not investigate how the multimedia can be controlled with gestures, but rather how the task of semi-autonomous driving can be executed with a gestural interface. Further along gestural interaction, we will study how people would use multi-touch gestures similar to the studies by Nielsen et al. (2003), Morris, Wobbrock and Wilson (2009 ; 2010), Hinrichs and Carpendale (2011). The result from these studies will be used as support for the process of designing gestures. Moreover, we will also explore gesture hinting proposed by Lundgren and Hjulström (2011). Lastly, we aim to investigate whether the system we will design will require a learning phase similar to the findings of Butz, Kramer and Rümmelin's study from 2013.

3. Theory

This chapter presents relevant theory for this study. It is divided in two main parts, affordances and theoretical framework for gestures. The first part describes the subject of affordances which includes mental models, conceptual models, and mapping. The second part presents theory within the domain of gestures. However, since there are no theoretical framework for the subject of gestures, we have compiled one based on the literature we have reviewed.

3.1. Affordances

Affordances are relationships between the properties of physical objects and the people examining them (Gibson 1979; Sharp, Rogers and Preece 2007; Norman 2013). They can be thought of as clues of how an object can be interacted with. For instance, a chair affords sitting and a glass affords drinking. Affordances exists naturally, even if they are not visible or known by the user. However, there is a difference between *real* affordances and *perceived* affordances (Norman 1999). The latter one refers to what actions the user think is possible, whereas the first one refers what possible actions there actually are. A fit between real and perceived affordances are of outmost importance in design work. The better the fit, the easier it will be for the user to understand and interact with the system.

Proceeding to mental models and conceptual models, these are two representation models that will affect the usability of a product or interface which designers need to consider in the design process. A mental model refers to how a user think a system works. It provides a deep understanding of how users think and operate when using a system (Lorenz, Kerschbaum and Schumann 2014). Furthermore, it can help designers make good decisions and ensure that the design make sense to the users. However, one must note that mental models are inaccessible for designers in an objective manner (Johnson and Henderson 2002) as they are are internal representations of a system (Greca and Moreira 2000). Moreover, mental models are incomplete and users often fail to remember every detail in their mental model (Norman 1983).

In contrary to mental models, conceptual models can be shared openly (Greca and Moreira 2000) as they are external representations of a system. A conceptual model refers to how a system is designed to work, and facilitates the understanding and teaching of a system. According to Johnson and Henderson (2002 p. 26) a conceptual model is: *"A high-level description of how a system is organized and operates."*. Further, Norman (2013 para. 1) states that *"A conceptual model is an explanation, usually highly simplified, of how something works"*. A well designed conceptual model allows users to understand how the system works (Norman 2013). It enables them to realize what will happen when they perform a certain action. However, it is important to note that the conceptual model is not about the look and feel of an interface (Johnson and Henderson 2002). It does not include commands, navigation, layout, data presentation etc. Neither is it equal to the user's' mental model of the system.

Mapping can be referred to as the relationship between elements of two sets (Encyclopedia Britannica 2015), e.g. a set of gestures and a set of functions. Thus, mapping is the concatenation of a set ofcontrols and the objects that they control (Norman 2013) and according to Cooper et al. (2014) mapping describes how a control, the object it affect and the intended action relate to each other. Furthermore, it is essential to consider the spatial mapping when mapping a set of controls as it will facilitate the user's understanding of the system (Norman 2013). For instance, the mapping of turning a steering wheel to the left to making the vehicle turn left facilitates the user's understanding of the vehicle's behaviour. It is important that the mapping relates to the user's mental model, otherwise the user will have to stop and try to understand how the system works. Moreover, it can break the user's flow, increase the cognitive load, and potentially lead to user errors if the conceptual model does not match the mental model.

3.2. Theoretical Framework for Gestures

Gestures can be divided into two parts, touch-based and free-form gestures, or two-dimensional and three-dimensional as they are also called. Zhai et al. (2011 p. 104) defines gestures (two-dimensional), or stroke gestures as they name it, to be: *"the movement trajectories of a user's finger or stylus contact points with a touch sensitive surface"*. A stroke gesture can be performed with either a bare finger or a tool, like a pen (also known as stylus) and is usually encoded as sequential movement in the x- and y-plane.

Further along the subject of gestures, Zhai et al. (2011) reasons about analogue and abstract gestures, where an analog gesture is one that try to mimic the physical world, similar to the spatial mapping discussed by Norman (2013). For instance, a swipe gesture could cause a document to pan, in the same manner it would in the physical world. Abstract gestures on the other hand, does not aim to resemble the physical world (Zhai et al. 2011). Such a gesture could be drawing an X in order to close a document, mimicking the icon for closing a document. However, Zhai et al. (2011) does not define analogue and abstract gestures as being one and not the other, but rather as a spectrum in which gestures can be both analogue and abstract. Gestures can be designed to partly resemble physical actions and partly the user's previous experience and conventions (Zhai et al. 2011; Saffer 2008; Andersson and Wikander 2013).

In addition to the properties of analog and abstract, gestures also have different attributes which varies depending on how the gesture is performed (Saffer 2008). For example, gestures can have different *Duration*, which refers to if it is a long or a short gesture. Further, the *Position* can vary as well as the *Motion* (direction and velocity of the gesture). The comptlete set of attributes is presented in Table 1.

Attribute	**Description**
Presence	Basic attribute, there must be something performing a gesture in order to trigger an action.
Duration	The duration of a gesture, eg. if it is short or long. Measured from the first touch to the last.
Position	The location of the gesture. Often determined using the location on the x/y axis but can sometimes be relative to the person performing the gesture.
Motion	The velocity and direction of a gesture. Sometimes actions are triggered by the mere presence of motion.
Pressure	Different pressure might trigger different actions in a system. Can be faked by measuring the size of the finger pad.
Size	The size of gestures can vary and depending on it additional input devices may be considered.
Orientation	The orientation of the device, or the person holding it.

Attribute	Description
Including objects	In addition to gestures, physical objects can sometimes be used to interact with the system. Interaction can be enhanced if the system is capable of recognizing objects in the user's hand and the type of the object.
Number of touch points	By tracking multiple touch-points different actions can be triggered if several gestures are performed simultaneously compared to if the gestures are performed one at a time.
Sequence	Gestures could also be performed in a sequence to trigger different actions than if performed separately. This demands systems that have several states, which is more complex for the user.
Number of participants	In some systems the number of participants should be considered as two persons interacting at the same time can be very different from one person interacting two hands.

Table 1. Descriptions of gesture attributes.

Saffer (2008) also define important characteristics for gestural interfaces. The first characteristic defined by Saffer is *Discoverable*, which means that it should be easy to understand what interactions that are possible. If the interface does not support this, it will lack the important affordances which is one of the main problems with gestural interfaces (Norman 2010; Norman and Wadia 2013). Furthermore, Norman and Wadia (2013) argues that gestural interfaces should provide immediate feedback which also is a characteristic defined by Saffer (2008) as *Responsive*. The complete set of characteristics are presented in Table 2.

Characteristic	Description
Discoverable	A major issue of gestural interfaces is the lack of affordances. Gestural interfaces does not have visible clues telling the user what parts of the interface that are interactive.
Trustworthy	The interface should give the user the impression that it is safe and trustworthy. Otherwise, the user will be hesitant to use it.
Responsive	High responsiveness is very important for gestural interfaces, since people are used to interact with physical objects and get instant feedback.
Appropriate	Gestures should be appropriate according to the culture and context they are going to be used in. Some gestures can be offensive in some cultures. Yet, that regards free form gestures.
Meaningful	Gestures should be designed so that they have a meaning for the person using them.

Characteristic	Description
Smart	Gestural interfaces should be smart and help the users with accomplishing the task at hand faster and more efficiently than without the system.
Clever	The interface should predict the needs of the user and adapt accordingly, since it will improve the user's experience.
Playful	A playful interface will encourage the user to explore it. Therefore, it should be designed to avoid errors and if it occurs, there should be an easy way to undo so that users never feel lost.
Pleasurable	Humans are more forgiving to things that are beautiful, therefore gestural interfaces should be designed to please the user both aesthetically and functionally.
Good	Design gestures that do not make the user feel embarrassed if performing them in public and design gestures that are easy to perform for people of all ages and health.

Table 2. Characteristics of a good gestural interface.

Despite the fact that the attributes and characteristics for gestural interfaces were presented by Saffer in 2008, Norman and Nielsen (2010) still claimed that there is a lack of standards for gestural interfaces. According to them gestures lack the essential affordances needed for successful human computer interaction (HCI) since they are ephemeral, i.e. there's no record of their path. Furthermore, Norman (2010) argues that if a system would be based only on gestural interaction, it would impair the user's possibilities to discover which interactions that are available. The main problems users have when using gestures is understanding what they can do, how and where they can perform gestures, what happens when they do a gesture, and how they can go back or undo what they did (Norman and Wadia 2013).

Even though there seem to be a lack of standards, there are a few guidelines on how to design gestures. In 2013, Anderson and Wikander presented a set of guidelines for gesture design in a vehicle context. These guidelines are based on research of gestural interaction for in-vehicle infotainment systems. A few of the guidelines are in line with the statements by Norman (2010) and Norman and Wadia (2013). For instance, the guideline *Use conventions* can be considered as counteract to the lack of standards since it advises designers to make use of already known gestures instead of inventing new ones. Further, the guideline *Ask users first* is in line with the characteristic *Meaningful* defined by Saffer (2008) as by asking users for input when designing new gestures will facilitate creating gestures that have meaning for them.

Guideline	Description
Use conventions	Make use of gestures that the user already know instead of inventing new ones.
Ask users first	Do not design new gestures without asking users for input, since people think differently. Adapt the gestures to known conventions, or follow the mental model of the majority of the users.

Guideline	Description
Use a sufficiently sized interactive area	Touch targets should have a minimum size of 1.5 x 1.5 cm. Yet, recommended smallest size is 2 x 2 cm.
Use patterns	Use patterns for the gestures. For example, a certain number of fingers could correspond to a certain category or function.
Common gestures should be simple	Commonly used functions should have simple gestures. Further, consider if the function is time critical or not. If so, make the gesture fast and simple.
Be forgiving	Make the system forgiving to the user as different persons perform the same gestures a little bit different.
Easy undo	Users are likely to make mistakes. Therefore, make it easy for the user to undo the last action.
Limit the number of gestures	The human working memory can only manage between 5 and 9 objects simultaneously. Hence, limit the number of functions that are interactable with gestures.
Provide multiple alternatives for interaction	Provide the users with multiple ways of performing the same action.
Allow the user to evolve into an advanced user	Develop several sets of gestures that are of different levels of complexity in order for the user to be able to first use the most simple gestures and later on start interacting with the more complex gestures.

Table 3. Guidelines for designing gestural interfaces.

4. Methodology

This section presents the research approach as well as descriptions of the different methods and techniques which we employed in the study.

4.1. Research Approach

In design work, there are different approaches which can applied in the design process. Examples of approaches are Human Centered Design (HCD), Activity Centered Design (ACD) and User Centered Design (UCD). The first one puts the human in the center and focuses on understanding the human; how they behave, what they want and need in their lives (Brown 2008). Moreover, HCD put focus on making products desirable for the users as well as the financial aspects of introducing the product to the market. Such a perspective would not be relevant for our study as it lies outside the scope. The second approach (ACD) focuses more on the activity rather than the human (Norman 2006). Since the scope of this study focus on designing an interface for semi-autonomous vehicles, it includes future products that probably not will be available on consumer market for many years. It would be difficult to study the activity of semi-autonomous driving. Therefore, it would not be suitable to base the research approach on ACD.

With this in mind, we decided to base the research approach on UCD. This allows us to focus on the interaction between the driver and the vehicle without a broader perspective questioning the human need of transportation, the need of autonomous vehicles or the role of the driver. Further, since there are little standards for designing gestures the study will require that users are strongly involved in the design process, which is central in UCD (Gulliksen, Lanz and Boivie 1999; Normark 2014). Involving users will facilitate making the design usable and suited for the users (Normark 2014). In addition, it simplifies the optimization of the HMI to be as fast and efficient as possible (Redström 2006). Lastly, it is important to observe users in a true context (Gulliksen, Lanz and Boivie 1999) since it will provide a deep understanding of the users and their behaviour.

An informal definition of UCD has been proposed by Gould (1988) and it consists of four basic principles of UCD, cited from Normark (2014 p. 44):

- *"Early focus on users. Have direct contact with users to understand their characteristics and the characteristics of their tasks."*
- *"Integrated design. All aspects of the system should be designed in accordance with each other under one management."*
- *"Early and continual user testing. Empirical data is needed to motivate design changes."*
- *"Iterative design. Modify the design according to the user testing."*

However, even though users should be strongly involved one must keep in mind that it is not the users that are supposed to produce the final design (Williams 2009). Moreover, when employing UCD, one should not ask the users directly what they want, but instead try to profile different users and their behavior. Asking them directly may lead to a bias understanding of the user.

4.2.1. Interviews

A technique for getting a thorough view into the users behaviour is to perform interviews (Toolkit, HCD 2008), which can be structured in four different ways; structured, semi-structured, unstructured and group interviews (Sharp, Rogers and Preece 2007). The approach of the interview should depend on the goal of it and if the goal for example is to get feedback on a specific feature of a design a structured interview is more suitable (Sharp, Rogers and Preece 2007). This kind of approach is similar to a questionnaire and the typically closed questions are used which means that there are a predefined set of answers to the question. The interviews should preferably be performed in the workplace of the users which will make the users feel more at ease (Toolkit, HCD 2008). Additionally it will be easier to understand users explanations when you can see the environment and the objects that they mention. A problem with interviews is that what users say they do does not always correspond to what they actually do, often they answer according to what they believe will show them in the best light (Sharp, Rogers and Preece 2007).

4.2.2. Think aloud

A method for understanding users and their impressions of an interface is the think aloud technique in which test subjects are asked to speak their thoughts out loud (Zhao and McDonald 2010). However, questions have been raised about the validity of the data from such studies because of interventions from the practitioner. Interventions are not part of the original method but have become common and are used for explanations alternatively to lessen the feeling of anxiety of the user. The risks of intervening during a test is that the participant might be helped in regards of completing the task at hand.

4.2.3. Brainstorming

According to Barker (1997), brainstorming was invented by an american working with marketing in the 1930's named Alex Osborn because he had discovered a pattern regarding new ideas in meetings. The pattern was that new ideas often were criticized and subsequently "destroyed" during the meetings. In response to this Osborn developed four guidelines for brainstorming sessions; no criticism, wild ideas are appreciated, as is quantity, and that one should try to evolve and combine ideas (Barker 1997; Wilson 2013).

Generating numerous ideas and avoiding criticism is also stated by Norman (2013) and Fullerton (2014). Furthermore, Fullerton (2014) states that brainstorming should take place in a creative environment (such as a dedicated room for brainstorming) since the normal workplace often can be rather uncreative. He also points at the importance of visualizing the ideas, for instance on a whiteboard or a larger piece of paper. By doing this, the ideas are shared to the whole group and it can lead to that new ideas are created as well as facilitate collaboration.

4.2.4. Low Fidelity Prototyping

Low fidelity prototypes, or lo-fi prototypes, are used in the process of generating and evaluating different design ideas (Benyon 2010). According to Benyon (2010), they are an essential tool when involving users in the design process. Lo-fi prototypes focus on the key functionality and overall design, and not so much on detailed design parameters (Saffer 2008). They are quickly made and suitable to use in the early stages of a project since they do not require a lot of time and are cheap (Sauer and Sonderegger 2009). Furthermore, lo-fi prototypes provides designers with a tool for evaluating different interaction approaches early in the design process (Virzi 1989).

4.2.5. High Fidelity Prototyping

High fidelity prototypes are similar to lo-fi prototypes but are more refined and closer to the final product (Saffer 2008). Such a prototype could be a molded piece of plastic with a dummy keyboard Sharp, Rogers and Preece (2007). Some of the problems with high fidelity prototypes are that they can take too long time to implement and a bug can stop the testing completely. On the other hand, the advantages are that they can be used for marketing purposes, and can have the look and feel of the product.

4.2.6. Participatory Design

Participatory design is a method in which users are involved in the design process in order to be able to influence the design of the product or service that they are going to use (Yamauchi 2012). Further, according to Yamauchi (2012) the basis of the method is in workshops and design sessions in which users are included. When included in the design process, users are able to propose their own solutions to the design. Gould (1988) argues that participatory design has its strength in that users doing the actual work with your product later on know much more about the work than *"drop-in designers"* (Gould 1988 p. 100). Moreover Gould (1988) also states that besides leading to better systems, the pride and self esteem can be increased and perhaps even increase the productivity.

4.2.7. Participatory Evaluation

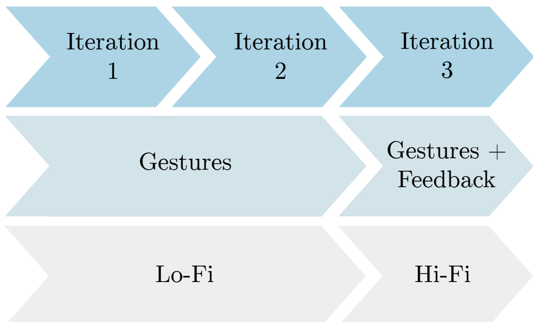
A method for evaluating an early design is Participatory Evaluation which is according to Maguire (2001) is an informal method and subsequently cost effective. Employing participatory evaluation is suitable for finding out the impressions users have of a particular screen of a design or what results they think that performing a certain action will yield. The method is employed by letting users perform a number of tasks whilst explaining what they do to moderator observing them. If the user is silent the moderator prompts the user for information about their actions.

4.2.8. Scenarios

A technique for describing how a user might interact with a system is Scenarios (Maguire 2001; Rosson and Carroll 2002; Sharp, Rogers and Preece 2007). They describe the setting in which the scenario takes place, the involved users as well as their actions and experiences of the system (Rosson and Carroll 2002). They can be of help when preparing for user evaluation studies, to clarify user requirements and help designers to think of the characteristics of the users and the context of use (Maguire 2001). Further, scenarios should be formulated in a language easily understood by stakeholders in order to be able to involve them in the design process (Sharp, Rogers and Preece 2007).

5. Process

The process was carried out in three iterations (see Fig. 2). In the first two iterations we focused on the design of gestures for the predefined use cases (see Table 4). Note that the use case stop does not mean that the vehicle will stop where it is. Instead, the vehicle will stop safely at the next possible parking space along the road. We involved users in design sessions with lo-fi paper prototypes on which they could perform gestures. Since we did not include any feedback it was enough with very simple prototypes for gathering data of what gestures that would most natural for the users. In the third iteration we moved to designing feedback and evaluation of the complete system. By doing that, we went from lo-fi prototypes to a hi-fi prototype with feedback in two modalities; visual and auditory. Overall, we involved users throughout the process since we considered it to be of significant importance since our task was to design a completely new system rather than improving a previously designed system.



Fiture 2. The process of this study divided into iterations, focus and prototyping fidelity.

Increase speed	Take first exit/turn
Decrease speed	Take second exit/turn
Change lane	Take third exit/turn
Move inside lane	Stop
Overtake	Undo

Table 4. Predefined use cases for the study.

5.1. Iteration 1

In the first iteration the goal was to design a first draft of gestures and evaluate the draft with users. Furthermore, we also aimed at collecting data about which gestures the users perceive as the most natural for each use case. In order to do that, we let users design gestures themselves in a vehicle context. This gave us a rich set of data which helped us understand which gestures the users considered natura. An overview of the first iteration can be seen in Table 5.

Goals	How	Key findings
• Design first draft of gestures. • Evaluate gestures with users. • Gather data about what gestures that are most natural for the users.	• Brainstorming gesture ideas. • Let users evaluate gestures. • Let users design gestures in vehicle context.	• Users performed mostly analog gestures (i.e. similar to the movement of the car). • Some users had difficulties designing gestures for second and third exit. • Users seemed to have a system thinking, i.e. wanted the gestures to be consistent and logical.

Table 5. Overview of iteration 1.

5.1.1. Methods and Design

We began the iteration with brainstorming gesture ideas, based on the use cases in Table 4. First, we did individual brainstorming and sketching of possible gestures. Then we presented the ideas to each other and began discussing them. The reason for starting individually was that we did not want be influenced by each other at first since that could have affected the diversity of the ideas. Some of the first sketches we did is presented in Fig. 3.

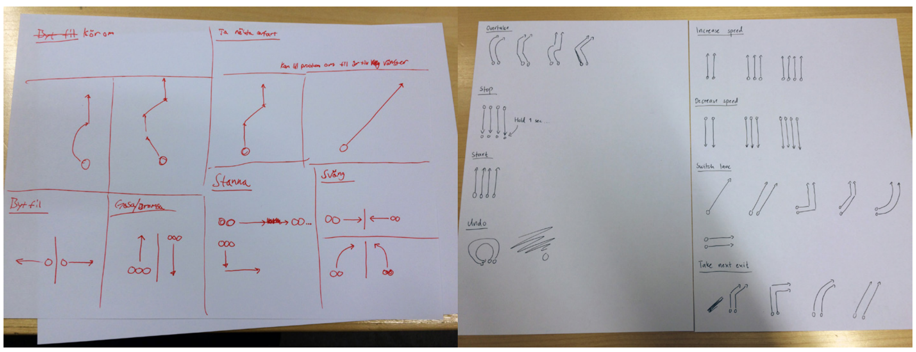


Figure 3. Early ideas of possible gestures.

After the ideation we conducted a first evaluation on our own, eliminating gestures that seemed too complex, unnatural or unintuitive. In other words, we selected the gestures that were perceived most natural and intuitive for us. These were mostly analog gestures (Zhai et al. 2011), which aims to resemble the actual movement of the vehicle. This became the basis for the for the upcoming gesture evaluation with users. However, since we did not have any similar system to analyze we decided to let users design gestures as well. To only evaluate the gestures we designed would not have been enough since we then might have omitted gestures that users could have thought of whilst we did not. Furthermore, as Norman and Nielsen (2010) stated, there are no standards for how to design natural gestures which makes it even more important to involve users in the design process. Additionally, by letting the users design gestures we would be more likely to find gestures that are natural and intuitive.

In the evaluation we decided to only evaluate single finger gestures (see Fig. 4) since including variations of each gestures with multiple fingers would create a too large set of different gestures. It would have been overwhelming for the participants to evaluate such a high amount of gestures. Furthermore, previous studies that we had reviewed (Morris, Wobbrock and Wilson 2010) indicated that users generally prefer single finger gestures. Therefore, we decided to exclude the use cases including take second/third exit/turn from this particular evaluation since those would require gestures of more complex nature and might have required multi-touch.

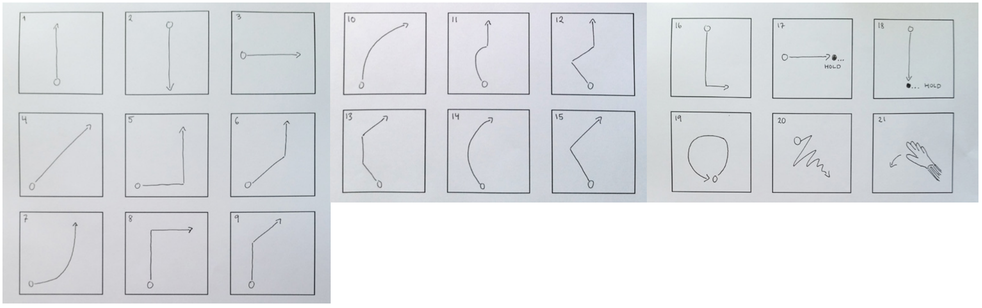


Figure 4. The gestures that the participants were supposed to map to the different use cases.

5.1.1.1. First Gesture Design Session

The first gesture design session with users included both participatory design and evaluation of gestures in the form of mapping gestures (see Fig. 4) to the different use cases. The participatory design was conducted in a vehicle context (see Fig. 5) and the evaluation in an office/home environment. For this session we involved 15 users in total.

The session began with an introduction to semi-autonomous driving and how the system was envisioned. We explained that the user would give the vehicle instructions by performing gestures on the input device, which then will be executed by the vehicle. Further, we told the participant that they would get feedback after inputting the instruction. However, we did not specify the design of the feedback since we at this point in the design process focused only on the gestures.

Figure 5. The three alternative positions for the input device in the first gesture design session.

After the short introduction, the participants were asked to choose between three different sizes of the input device (12x16 cm, 15x20 cm, 18x25 cm). For this we used simple paper prototypes of cardboard and let the participants choose which size they would prefer. Then we asked them where they would prefer to have the the input device; on the steering wheel, on the center stack, or between the gear stick and the armrest.

The next step was to let the participants design gestures based on a scenario (see Appendix A), which included all the predefined use cases. This time we decided to include second or third exit/turn since we wanted to investigate which gestures the users would design for those use cases. The gestures were performed on the paper prototype with the size and placement of their choice. The moderator read the scenario out loud (one use case at the time) after which the participant was asked to perform the gesture he or she thought was most natural. During the interview the participants were encouraged to think aloud and explain the design of their gestures. The interview was recorded with a smartphone taped onto the rear view mirror of the car capturing the participant's gestures and voice. Additionally, the moderator also took notes on the gestures.

After the design session the participants were asked to map the gestures that we had designed to the different use cases (see Fig. 6). Furthermore, they were also asked to rank the gestures within each use case so that we could see which shape they found most natural. However, as mentioned before the use cases take the second/third exit/turn or turn was not included in this evaluation.

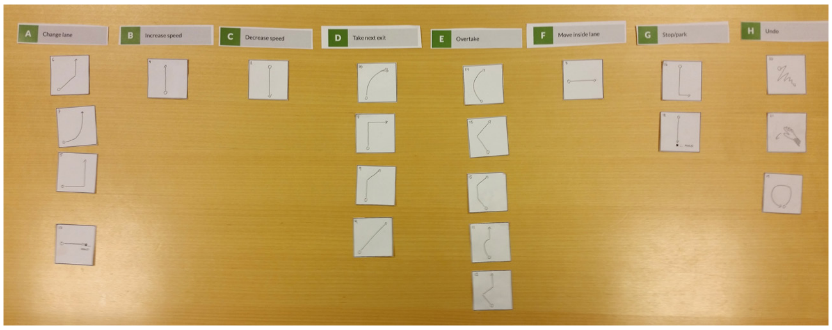


Figure 6. An example of how it could look when a participant had finished the mapping and ranking of the gestures.

5.1.2. Findings

Patterns from the participatory design session are presented in Fig. 7. Each gesture presented was performed by at least two users. The gestures in green were performed by most users. The gestures with blue color were the second most common category and the gestures in purple was the third most common ones. The result indicates that users prefer single finger gesture in general, similar to the findings of Morris, Wobbrock and Wilson's (2010) study. However, there were also participants who prefered multi-touch. For instance, one participzant stated that: *"Using two fingers is more distinct and clear"*. Furthermore, we could see that the gesture for increase and decrease speed seemed to be very natural since 14 of 15 participants performed a swipe upwards for increase speed and a swipe downwards for decrease speed. Moreover, a swipe to the left/right also seemed to be natural for change lane since several participants performed this gesture and did it without reflecting too much.

Regarding the use case take next exit/turn there were a larger variation of gestures than for increase/ decrease speed and change lane. It did not seem to be as straightforward to design a gesture for this use case as for the previous use cases. Continuing to the use cases for second and third exit/turn, the variation was even larger. Here, there were no gesture which was more common than any other. Instead, the patterns we could see was four different gestures, which was performed by two participants each. We could see that the participants took more designing these gestures. They seemed to have more difficulties choosing a gesture for second and third exit/turn, which indicates that designing gestures for these use cases are not as natural as for the other use cases. Moreover, when the users got to the use case take the third exit/turn, they stopped for a while and realized that there should be a pattern for the gestures across first, second and third exit/turn. Due to that, several participants changed the previous gestures for first and second exit/turn so that it got consistent pattern. In similarity, we could see that the participants had some difficulties designing a natural gesture for the use case stop since they needed to spend more time on the design than for some of the other use cases. Furthermore, the variation of gestures was quite noticeable. Only two participants performed the same gesture which could indicate that there is no natural gesture for this use case.

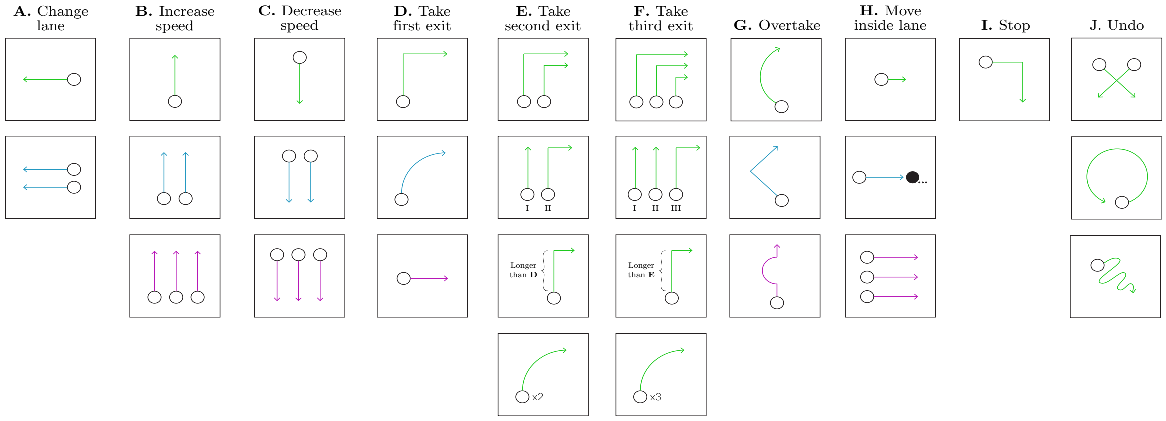


Figure 7. Patterns in the participants' gesture design. The gestures in green where most common, the blue were second most common and the purple were third most common.

Continuing to the use case overtake, all the participants performed analog gestures which aimed to resemble the movement of the car. The different gestures performed by the participants can be seen as variations of the same gesture which is similar to findings by Hinrichs and Carpendale (2011). However, it seemed as the simpler the gesture motion was, the easier it would be to perform. During the gesture mapping of this use case one participant stated that: *"The fewer steps the better"* by which he meant that gestures with less direction changes were easier to perform (see Fig. 8). Lastly, for the use case undo there were three patterns but similar to take the second and third exit/turn, there were no gesture that was more common than another.

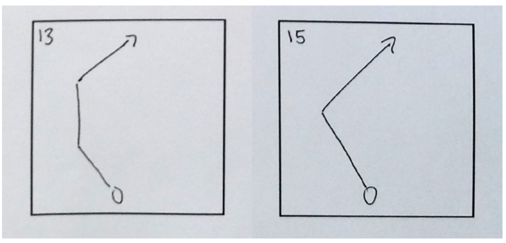


Figure 8. Two gestures where the right one has fewer steps than the left one.

Regarding the choice of placement for the input device it was fairly widespread across the group. One user that wanted the input device in the steering wheel expressed that: *"I want all the steering in one place"* and another user that preferred the center stack stated that: *"The center stack is where I'm used to interact"*. Additionally, a user that preferred the position behind the gearstick rather than the center stack commented that *"The center stack is too far away"*. All in all, seven users prefered the placement behind the gearstick, four wanted it on the steering wheel and four prefered to have it on the center stack. This points towards that users prefer the placement behind the gear stick, but it cannot be generalized due to the fairly small size of the test group. Moreover, there are also other arguments for placing the input device behind the gear stick. Having the input device as part of the steering wheel would require the center of the steering wheel to be fixed since a rotating input device since would impair performing the gestures. Therefore, we did not consider the steering wheel to be a suitable placement. Neither did we regard the center stack as suitable since it already contains an interface. Furthermore, several users mentioned that the center stack is too far away for performing gestures. Therefore, we decided to place it between the armrest and the gear stick. We would argue that this placement would be suitable since this area often is used for cup holders.

5.2. Iteration 2

The second iteration continued on the same track as the first iteration with a strong focus on the gestures. The goal of this iteration was to determine a set of gestures for the system. In order to do that we designed two gesture concepts for each use case based on the findings from the previous iteration. These concepts were then evaluated with users. Additionally, we let the users design gestures once more but with the input device positioned at the same place. An overview of this iteration is presented in Table 6.

Goals	**How**	**Key findings**
• Determine a set of gestures for the system.	• Let users design gestures with the input device placed at the same position. • Let users evaluate gesture concepts.	• Users seemed to prefer single finger gestures. • Users seemed to prefer 2 and 3 fingers rather than length for taking 2nd/3rd exit/turn. • Users seemed to prefer straight gestures rather than curved.

Table 6. Overview of iteration 2.

5.2.1. Methods and Design

Based on the patterns from the first iteration we designed two gesture concepts. However, we did not only look at the patterns since we had realized the importance of having consistency across the whole set of gestures, in line with the guideline *Use patterns* (Andersson and Wikander 2013). Therefore, our design originated from the most common gestures in the previous gesture session, but including some changes in order to reach consistency. The two concepts are presented in Fig. 10.

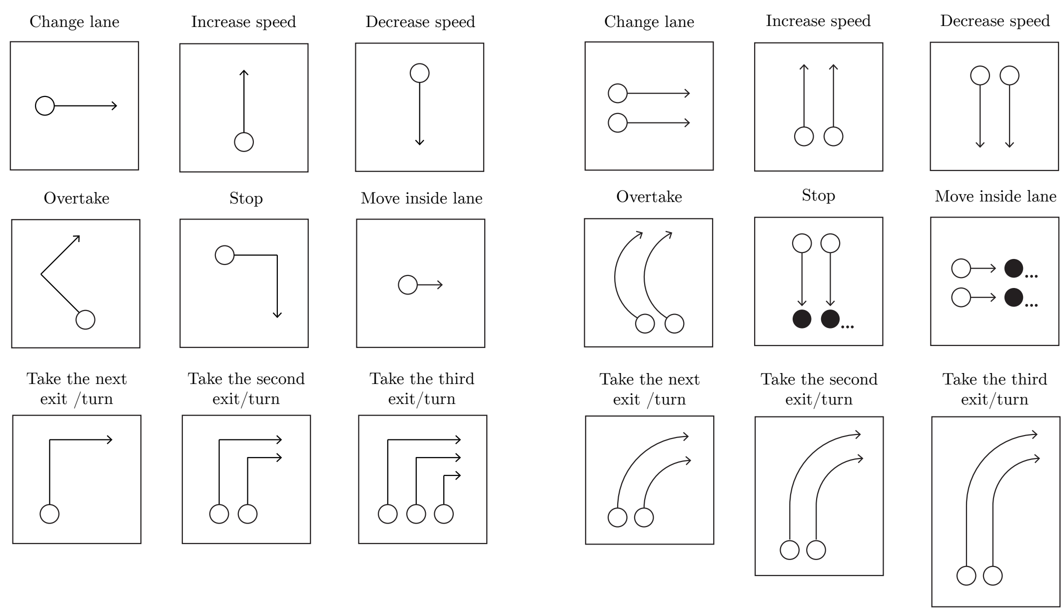


Figure 10. The two gesture concepts that were evaluated with users. To the left is concept 1 and to the right is concept 2.

The first gesture concept was based on using one finger for each use case, except for the second and third exit/turn in which multi-touch was used. The second concept was similar to the first but used two fingers for all gestures instead of one and had the length of the gesture as indicator for first, second or third exit/turn. Furthermore, some of the gestures were curved rather than straight as in concept 1. Note that the gestures for undo (see Fig. 11) was evaluated separately since we had three different gestures which we wanted to evaluate for this use case.

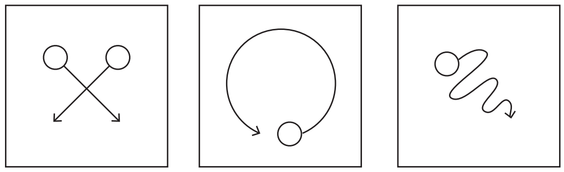


Figure 11. The three gestures for the use case undo. Note that the first gesture is performed in two seperate motions.

5.2.1.1. Second Gesture Design Session

The second gesture design session was conducted in the same environment as previous one in the first iteration, and with the same 15 users except for one which was unavailable and therefore got replaced by another participant. The input device was placed between the armrest and gear stick for all participants. The procedure was the same as in the previous gesture design session, but this time the order of the uses cases in the scenario (see Appendix B) was different. Furthermore, instead of the gesture mapping we let the users evaluate the two concepts which we had designed. This was done by going through the scenario once more and for each use case the participant got to choose which concept had the best gesture for that particular use case (see Fig. 12). Yet, if they thought none of the concepts had a natural gesture, they were also allowed to design their own. Similar to the previous session the participants were encouraged to think aloud and motivate their choices.

Figure 12. A participant discussing with the moderator regarding the use case decrease speed.

5.2.2. Findings

Starting with the results from the participants gesture design, the results were similar to the previous gesture design session. It seemed as the placement did not have significant impact on how the participants performed their gestures. Most participants performed single finger gestures, which is similar to the findings of Morris, Wobbrock and Wilson (2010). However, several of the participants also expressed that it does not really matter, which is in line with the findings of Morris, Wobbrock and Wilson's previous study in 2009. Patterns from the second gestures design session is presented in Fig. 13.

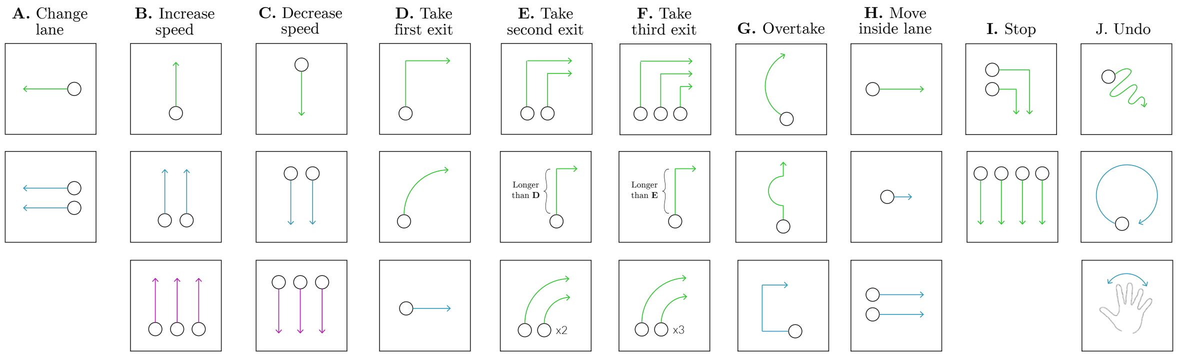


Figure 13. Patterns in the participants' second gesture design. The gestures in green where most common, the blue were second most common and the purple were third most common.

The result of the evaluation showed that most participants preferred concept 1 which indicates that straight gestures are generally more prefered than curved. Some users expressed this fact during the evaluation and there were just a few which prefered curved single finger gestures rather than straight. The complete result of the evaluation is presented in Table 8.

Use case	**Concept 1**	**Concept 2**	**Other**	
Increase speed	9	5	1	
Decrease speed	9	5	1	
Change lane	12	3	0	
Move inside lane	3	8	4	
Overtake	7	4	4	
Take first exit/turn	10	1	4	
Take second exit/turn	10	3	2	
Take third exit/turn	10	3	2	
Stop	4	8	3	
	Cross	**CCW circle**	**Eraser**	**Other**
Undo	2	4	8	1

Table 8. Result of the evaluation of the two gesture concepts and the gestures for undo. The numbers represents how many particpants that prefered the particular gesture.

When asked about their preferences for the use case take second or third exit/turn most of the participants preferred the first concept since they found it more intuitive to let the number of fingers determine which exit should be taken rather than the length of the gesture. However, that might have been influenced by the fact that some of the participants were used to using multi-touch gestures on laptop touchpads. Furthermore, most of the participants also had questions about if they would get any feedback when they had performed a long enough gesture for the second concept. That could indicate that feedback in the input device is needed for this concept to function. Continuing to the undo instruction most users preferred the erase motion. One user said that performing the counterclockwise circle (similar to the undo icon in e.g. Microsoft Word) could mean that you want to make a U-turn and drive back to where you came from.

Regarding the placement of the input device, the participants were very positive towards placing it between the gear stick and the armrest. They thought it was a natural position and comfortable to have the arm on the armrest while performing the gestures.

5.3. Iteration 3

In iteration three we changed focus from designing gestures to designing feedback. The goal of this iteration was to design the feedback and then evaluate the whole system with users. We started by brainstorming ideas for feedback, focusing on the visual modality. Later on we also conducted a workshop with other interaction designers in order to get more ideas. Thereafter, we designed gestural hinting which we evaluated with a couple of users. Lastly, we tested the whole system in a simulator. An overview of the iteration is presented in Table 9.

Goals	**How**	**Key findings**
• Design feedback for the system. • Evaluate gestural hinting. • Evaluate the whole system (both gestures and feedback).	• Brainstorm feedback ideas. • Test gestural hinting separately. • Test the concept in a simulator.	• Gesture hinting was not helpful and difficult to calibrate. • Some users had difficulties remembering the stop gesture. It seemed to be the least natural one. • The multimodal feedback was helpful. • Some users wanted the feedback to be more connected to the gestures. • Some users had difficulties understanding the symbol for unknown gestures.

Table 9. Overview of iteration 3.

5.3.1. Methods and Design

The ideation of the feedbacks started with a brainstorming session in which we ideated symbols for every use case (except undo). The symbols was supposed to provide information of which instruction the user have given to the vehicle. The session was conducted in a separate room where the ideas was shared openly and drawn onto a whiteboard (see Fig. 14). We continued until we had several symbols for each use case.

After having ideated several symbols we selected one symbol for each use case which we thought was the most informative and easily understood. For increase and decrease speed, we selected symbols which were similar to those used for the cruise control in cars since they could be more recognizable for users with previous driving experience. For the first, second and third exit/turn use cases we selected symbols that resembled road signs since those use cases regards navigation.

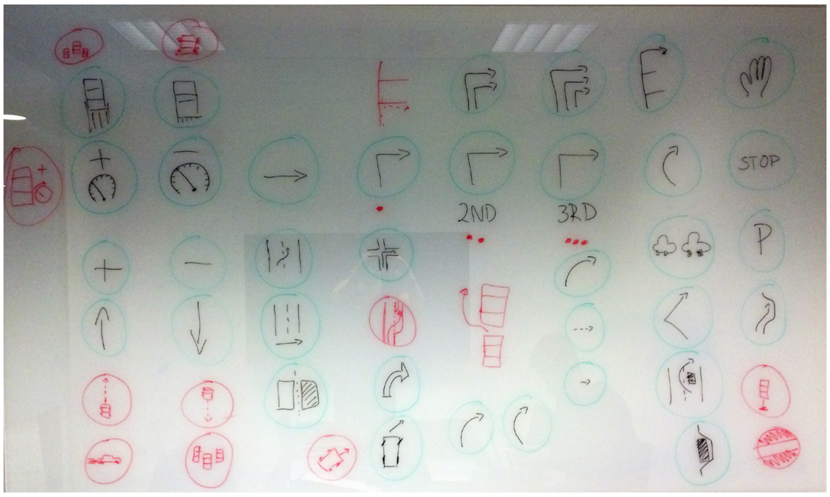


Figure 14. Early sketches of symbols for the different use cases.

For the auditory feedback we started with a workshop with a senior interaction designer at Volvo Cars who explained the basis for in-vehicle auditory feedback. That provided useful insights for us, which we could apply in the process of designing our own auditory feedback. The feedback consisted of three parts; a confirmative sound for when an instruction is confirmed, a rejecting sound if the gesture is not recognized by the system, and one sound for undo.

Regarding the use cases, we had a discussion with the stakeholders at Semcon and Volvo whether the move inside lane was useful or not. To us it seemed to be redundant and too much operational control. We discussed the matter with a the users who all stated that such a use case would probably be redundant since the vehicle should be driving in the middle of the lane at all times. However, there could be situations where the driver might want to move to the side a bit in order to facilitate an overtake on a highroad. Therefore we decided to change the use case move inside lane to move to the side and incorporate it with change lane. By doing that the user could perform the same gesture and the vehicle whether it should change lane or move to the side depending on if it is on a highway or a highroad.

Lastly, we decided to change the gesture for undo to be a double-tap instead of an eraser motion. The reason for that was that we considered the eraser motion require too much effort from the user. Furthermore, switching to the the double-tap gesture made it more differentiated from the other gestures.

5.3.1.1. Gesture Hinting

In order to provide further means of feedback to the user and increase the discoverability of a gestural interface, gesture hinting have been suggested by Lundgren and Hjulström (2011). Therefore, we designed gesture hinting for our input device. The hints worked so that whenever the user started gesturing, the interface would show hints for the gestures that still are possible. For instance, if the user would start moving the finger forward, the symbol for increase speed would show on the top, and on the sides the first exit/turn symbols would show. These symbols can be regarded as goals which the user should aim for in order to complete a gesture.

We tested the hinting with a the users and we could see that the hints was not providing much help for the users. They seemed to disregard the hints and performed gestures without making use of them. Furthermore, the users performed gestures of various sizes which made it very difficult to calibrate the position of the hints. An example of this was when users performed the take first exit/turn to the right gesture. In that case some users performed a quite long swipe upwards, which positioned the hint for completing this instruction to the lower right, thus making some users complete the gesture incorrectly, leading to an unrecognizable gesture. An example of can be seen in Fig. 15.

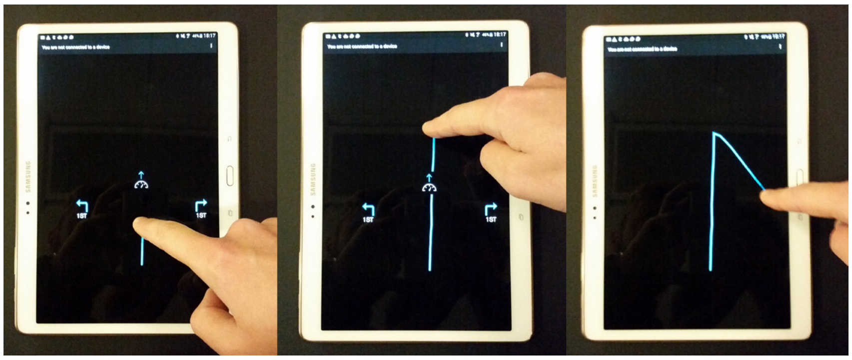


Figure 15. An example of gesture hinting. The user wants to perform the gesture for "take next exit/turn right" and starts with a swipe upwards. Unfortunately, the swipe is a bit long and the hint is positioned to the lower right twhich makes the user complete the gesture incorrectly.

5.3.1.2. Concept Description

Based on the findings from the tests of gesture hinting we decided to remove it from our concept. Yet, we kept the finger trail from the user's interactions since that helps the users to understand which gesture they have performed. Furthermore, Norman (2010) have previously pointed at the importance of providing a trace for gestures.

The gestures which were evaluated are presented in Fig. 16 and the visual feedback in Fig. 17. We based the gestures on the first concept since that was most preferred by the participants. The feedback was based on the symbols which we selected earlier in the iteration. The symbols with blue color are those that were shown when an instruction was confirmed by the vehicle and the one with a vehicle and a red arrow represented the case of rejected instruction. For instance, if the user would give the instruction to overtake when there is no vehicle to overtake it would be rejected by the system. Lastly, the red symbol with a counterclockwise arrow was shown whenever a user performed a gestures which was unrecognizable. It was also accompanied with the text "*Unknown gesture. Please try again*".

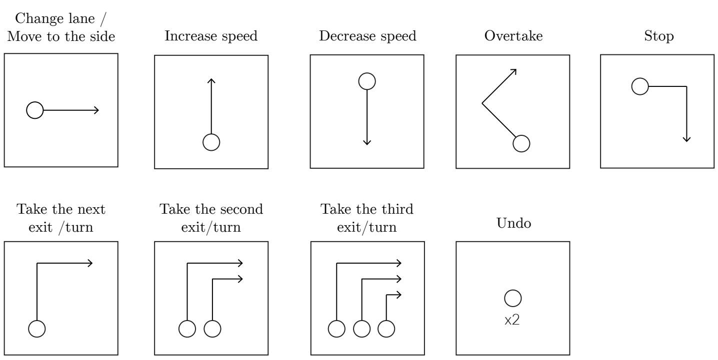


Figure 16. The complete set of gestures that the participants evaluated in the simulator test.

In addition to the visual feedback the concept also included auditory feedback. There were three sounds as mentioned earlier; one sound indicating that an instruction have been confirmed, one indicating that the instruction is not possible at the moment or that the performed gesture cannot be recognized, and one last sound confirming that the user have performed the undo gesture.

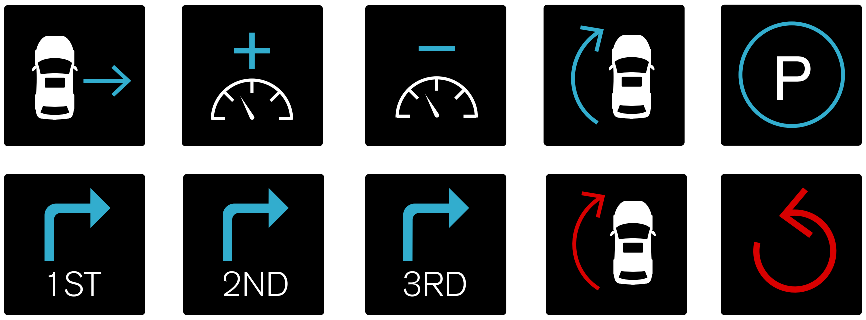


Figure 17. The visual feeback used in the simulator test.

5.3.1.3. Simulator Design

The concept we designed needed to be evaluated with users and we wanted to conduct it in the same context as the previous gesture design sessions but with an increased realism providing the feeling of actually travelling in a vehicle. Therefore, we decided to build our own high fidelity simulator. We did this with the use of a real car and a projector that projected a video on the wall in front of the car. The video was used to simulate the use cases we wanted to evaluate, such as change lane or overtake.

We recorded the video material in two attempts. In the first attempt we used a regular video camera mounted on a tripod in the front left seat of a right steered car. The tripod was fixed with the seatbelt so that it would not shake or move around during the drive (see Fig. 18). However, after reviewing the video material we realized that the video material was not as steady as we wanted. The video recordings was a bit shaky due to small vibrations inside the car. Furthermore, some parts of the interior of the car was visible in the video. We needed another solution that could record stable video without any shaking and preferably without visible interior. Therefore, in the second attempt, we decided to use a camera mounted on the hood of the car (see Fig. 18). With that we repeated the same procedure as before, recording every use case that we needed. This method gave us much better video material, not only was the material more steady but it was also in a wider format without any visible interior of the car.

Figure 18. The image to the left shows the camera mounted on a tripod inside the car. The image to the right shows the second attempt with a GoPro camera mounted on the hood.

However, the sound from the recordings were not as good as in the first attempt. We wanted the sound from the inside of the car and now we got the sound from the outside instead which mostly consisted of wind noise. Yet, we still had the previous recordings from the camera inside the car and from those we could take the in-vehicle sound and use it together with the new video recordings. We used one sound clip with a steady in-vehicle sound and adjusted the volume so that if the car would slow down in the video, the volume of the sound would also decrease in order to make it more realistic. The video material was then edited so that we had one short video clip for each use case that we wanted to include in the test.

On the roof of the car we placed a projector which projected the previously recorded video material on a white wall in front of the car, simulating that the car was moving along a road (see Fig. 19). Inside the car, we placed one tablet behind the steering wheel, acting as a DIM (Driver Information Module) where the driver would get feedback. Next to the driver's seat, between the gearstick and the armrest, we placed another tablet which was the input device for the gestures. In the back, behind the rear seat headrests we placed speakers simulating the in-vehicle sounds. They were placed face-down to make the sound a bit muffled and to be perceived less directed. In the front of the car, we placed two speakers for the auditory feedback. There we also placed a camera which recorded where the participant directed his or her eyes during the test. Lastly, we placed a smartphone under the rearview mirror, to record the participant's interactions with the input device.

Figure 19. The simulator we designed for the final evaluation. On top of the car is a projector placed to project the pre-recorded video on the wall in front of the car.

Each participant was seated on the driver's seat and in the passenger seat a moderator was seated, giving instructions to the participant throughout the test. The test was divided in two parts were the first part was testing our prototype and the second part was an interview about the usability of the system.

5.3.1.4. First Pilot Test

Before performing the actual user test we performed two short pilot tests. The first test was performed in a meeting room with the participant seated close to a TV screen with a table in front of the participant. A computer connected to the TV, which played the video footage of each use case. The input device was positioned flat on the table, close to the participant and the output tablet, which simulated a DIM, was placed at an angle a bit further away on the table. The complete setup is displayed in Fig. 20. In this pilot test we wanted to see if our test method would work and if the timing of the video clips worked well or not. It was important that the participant had enough time to perform the gesture for each video clip, but also that it did not take to long for the car to execute it. When the pre-recorded video were played and the participant were asked to perform a certain instruction, such as take the next exit right or change to the left lane. In the video clips, the time before the vehicle executed the instruction was approximately between 10 - 20 seconds depending on the use case.

Figure 20. The setup used for the first pilot test. The tablet flat on the table is the input device and the tablet at an angle is the feedback device.

The pilot test showed that this method seemed to work well. The participants were able to perform the gestures in time without any major difficulties. Even if they did the wrong gesture they could, with help from the moderator, perform the correct gesture before the instruction was executed in the video.

The participants had comments both regarding the test method itself as well as the design of the concept. One of the participants stated the he was not sure whether the video itself provided any real value to the test or not. Feedback on the system was that there should be different feedback for when an instruction has been confirmed and when it is being initiated by the vehicle. Other things mentioned was that they appreciated that the gestures could be performed a bit sloppy and still be confirmed and that the feedback for take the second or third exit/turn could have been more clear. In detail one participant thought that when a take second or third exit/turn instruction had been initiated and the car drove by an exit, not only the text displaying which exit the car is going to take should decrease but that there should also be some visual change to the symbol as well. Otherwise they thought that the symbols were good. Another comment was that it was good that the feedback symbols reminded of the shape of the gesture for respective instruction. The participants also thought that the gestures were intuitive and that the blue color used for various feedback was good, they thought it communicated that the system was “intelligent”. Lastly, one participant wished for voice feedback such as *“taking next exit”* when an instruction is confirmed.

5.3.1.5. Second Pilot Test

The second pilot test was conducted in the simulator described in section 5.3.1.3. with the participant seated in the driver’s seat. In this test we wanted to make sure that everything worked well and that the questionnaire would provide the data we needed to collect. Furthermore, since we had designed feedback for a rejected instruction and unknown gestures, we wanted to test that as well.

The result showed that the test worked well in general. There were no significant problems that needed to be addressed. However, the participants stated that the auditory feedback for when an instruction is confirmed and for when it is rejected was too similar and that it was a bit difficult to distinguish between them. Furthermore, the participant thought that the rejecting sound was a bit too positive. Therefore we lowered the pitch of it in order to make it more negative and therefore more distinguishable from the acceptance sound. Moreover, we also modified the scenario to include sections for testing the undo use case and the feedback for when the user attempts to give an instruction to the vehicle that cannot be performed, e.g. to overtake a nonexistent vehicle.

5.3.1.6. Final Evaluation

The final evaluation was conducted with the setup described in section 5.3.1.3. and with eight participants in total, six men and two women. Five of the participants had been previously involved in our study, and three had not seen the system before. We decided to go for eight participants since most of the usability issues is covered with five participants and more than ten participants will not provide more meaningful insights, but rather makes the data analyzing more complex (Lowdermilk 2013).

The test started with a short introduction of the context, semi autonomous vehicles and the basics of our concept. Thereafter, we explained the concept more in detail, both the gestures and the feedback. Furthermore, we also explained how to input instructions and where the feedback would be displayed. Further, the participant also got to test the system by performing a couple of gestures and getting the relevant feedback even though no video footage was played at this stage. When the participant had tried out the gestures and was familiar with the system, we began the test. It started with short video in which no instructions were performed, to get the participant in the mindset of traveling in a self driving car. After the test drive we continued by playing a number of video clips of the use cases of our concept, in total 8 minutes of video. In each subsequent video clip the participant was instructed to give a certain instruction to the vehicle, such as take the third exit to the right, change to the left lane or increase the speed. If a participant had trouble remembering a gesture for an instruction the moderator reminded the participant and the video clip was played once again. See Fig. 21 for an example of how it could look like during the test.

Figure 21. In this picture the participant have recently performed the take next exit/turn right gesture.

When each video clip had been played, the first part of the test was finished and the second part started. In this part we conducted an interview about the concept containing questions about how difficult the system was to use, what problems the participant had using the system, how easy the gestures were to perform, remember and how suitable they were. We also asked about the feedback; if they perceived it, if they understood it, what parts of it that they used and how they knew that the system had understood the instruction that they wanted the vehicle to perform. The questionnaire can be viewed in its entirety in Appendix C.

5.3.2. Findings

The result showed that the participants found the system easy to use. After the test we let the participants rate how difficult the system was to use on a scale of 1 - 10, where 1 was very easy and 10 was very difficult. All users rated the system between 1 and 3 and the average rating was 2.5 (see table 10). This indicates that the system was perceived as being easy to use. Furthermore, there was no significant difference between the five previous users and the three new users in terms of performance, which indicates that the system can be regarded as intuitive and that it has a fairly short learning phase.

	User 1	**User 2**	**User 3**	**User 4**	**User 5**	**User 6**	**User 7**	**User 8**
Rating	3	3	2	3	3	2	1	3

Table 10. Each particpant's rating of the system.

Regarding the gestures, the participants found them intuitive and clear. Some participants stated that the gestures were natural and logical. Others mentioned that the gestures was easy to remember. This was also something we could observe in the test since no user did more than one error despite only a few minutes of practice. A few statements by the participants during the test were: *"The gestures feel natural to me."* (User 4), *"The gestures are intuitive."* (User 6), *"The gestures are easy to remember."* (User 2).

However, the participants also had also some critical feedback about the gestures. Some of the participants had difficulties remembering the stop gesture. Three out of eight participants performed a wrong gesture for this instruction. Instead of swiping right first and then down, they did the opposite and swiped down first and then to the right. Other participants, who performed gesture correctly, also had to think a bit longer before performing this gesture. Furthermore, one of the participants that performed this gesture incorrectly, stated that:

> *"It feels more natural to start with swiping downwards and then to the right, since that is more similar to what the car actually do. Braking first, then moving to the side."* (User 7)

Additionally, one user performed the overtake gesture incorrectly. That participant stated that it did not feel as natural as the other gestures. Yet, she thought it would be easier after having used the system for a while. Regarding the number of fingers, the participants' thought it was easy to use one finger for the gestures. They would not prefer to have multi-touch gestures in general. However, they thought using multiple fingers for second and third exit/turn was easy and logical, since those are only variations of the gesture for taking the first exit/turn. One participant stated that:

> *"It was good that the gestures for 2nd and 3rd exit/turn was just a variation of the gesture for 1st exit/turn. That's easier to remember."* (User 2)

All of the participants' expressed positively towards the fact that we had changed the undo gesture since that makes it more differentiated from the other gestures. They found the new gesture easy to understand and perform. One participant stated that:

> *"The double-tap gesture for undo is easier and quicker to do than the previous erase gesture."* (User 2)

However, a few of the participants had concerns regarding the double-tap gesture since it in a computer context usually means confirming or selecting something. Yet, they did not think of it as a significant problem.

Moving from the gestures to the feedback, the participants seemed to find the multimodality helpful. The visual feedback provided them with information about what the vehicle would do next and the sound confirmed whether they had done right or wrong. Some of the participants expressed that they mostly used the visual feedback to know whether they had performed the correct gesture, whilst a couple of other participants relied more on the auditory feedback. Most of the participants thought the symbols were clear and easy to understand except the symbol for unknown gesture which was perceived as quite unclear and not that easily understood for several users. The problem with this symbol was that the participants thought it looked like an undo symbol instead of retry, which was our intention. Furthermore, they found it inappropriate to color the symbol red since that color communicates stop:

> *"The symbol for when I've done something wrong looks like undo. And it is red, which does not feel right when I'm supposed to do something".* (User 8)

Moreover, participants also expressed that it was a bit confusing that the symbol for unknown gesture had an arrow when it did not communicate the planned movements of the vehicle at all, in contrast to the symbols for successfully performed gestures. Additionally, one participant thought that the symbol for stop was confusing as he thought it meant that the car would begin searching for an empty parking spot at a parking lot. Lastly, two of the participants thought that the overtake symbol looked a bit like a rotating car instead of overtake the next car.

Other comments mentioned by the participants suggested that that the symbols and the textual feedback generally should be bigger since they're such an important part of the driving. Further, one participant expressed that the feedback should look like the gestures as it could facilitate learning and remembering the gestures. The participants' also appreciated that the feedback was instantaneous but several wanted it to be presented in a HUD (Head Up Display):

> *"I want the feedback to be more connected to the gestures."* (User 8)
> *"It was good that I got immediate feedback".* (User 6)
> *"I'm used to look at the road so I want the feedback displayed in the HUD".* (User 5)

The participants expressed positively towards the sound used to mediate that the system did not understand a gesture or that the gesture could not be performed e.g. overtake a nonexistent vehicle. Two of the participant stated that:

> *"It was clear that I've done something wrong, but it did not sound urgent". (User 1)*
> *"The sounds were efficient, I immediately understood whether I've done right or wrong".* (User 2)

Another participant expressed that the denying sound was helpful since it was clear that she had done something wrong but that the sound was not too harsh. Generally the participants' also thought that difference between the sounds for confirmed and rejected instruction was clear. However, one participant expressed that the sound for confirmed instructions was a bit unpleasant and could become annoying after a while of using the system:

> *"The confirming sound is a bit irritating and too high in pitch."* (User 6)

Regarding the visual feedback in the input device, most of the participants did not make much use of it while performing the gestures. They did not look closely at the input device, but rather glanced at it before performing a gesture. However, they found it helpful whenever they performed a gesture incorrectly since they then looked more carefully on the input device and which facilitated performing gesture.

During the evaluation we could see that most user performed the correct gesture but the system sometimes misinterpreted their gesture due to technical issues. That indicates the importance of having a forgiving system. This fact was also expressed by one of the participants during the evaluation. Moreover, it is in line with the guideline Be forgiving from Andersson and Wikander's (2013) study.

Another insight from the test was that it could be more clear what the vehicle is currently doing. One participant expressed that it is not clear who is responsible for the safety. For instance, when the instruction overtake is initiated and the vehicle is waiting for a safe opportunity, the feedback is not explicitly communicating that the action is pending or the reason for it.

6. Final Design

Based on the findings from the final user test we made a few modifications according to the participants input. Starting with the gestures, we decided to change the gesture for stop from a swipe to the right and then downwards to be a swipe down and then to the right instead (see Fig. 22). That makes it more related to the movement of the vehicle. Moreover, the shape of it becomes a bit more differentiated from the other gestures. The complete set of gestures in our final design is shown in Fig. 22. Note that the gesture for taking second or third exit/turn are not performed precisely as they are presented in figure 22. In order to perform these gestures, the user simply swipe upwards with two or three fingers and then to the right without turning the fingers according to the illustration.

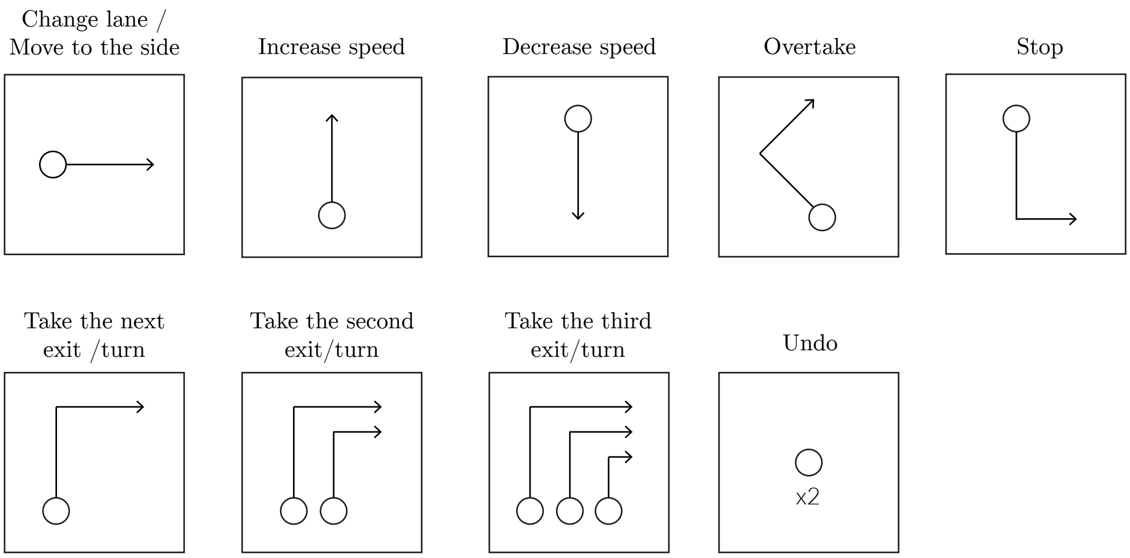


Figure 22. The complete set of gestures. Note that the undo gesture is a simple double-tap.

Regarding the feedback, we redesigned the icons for increase speed, decrease speed, stop and overtake to be more similar to the shape of the gestures for each instruction. By doing that, we have incorporated the shape of each gesture in every feedback symbol. For the use case increase speed, we changed the dial on the speedometer icon to be directed to the right in order to resemble higher speed as well as to make it a bit more differentiated from the decrease speed icon. Additionally, we redesigned the icon for unknown gesture to be a question mark instead as we would argue that it better communicates that the user have performed a gesture not recognized by the system. The set of feedback symbols is presented in Fig. 23. Each symbol correspond to a gesture, except the red question mark which is shown if the user performs a gesture which is not recognized by system. Note that the symbol for overtake with a red arrow is an example of the symbols for when an instruction is not possible to execute. In other words, it's the same symbol as for a confirmed instruction but with a red arrow instead of blue, accompanied with a text explaining why the instruction is not possible. Lastly, the pitch of the confirmation sound in the auditory feedback was lowered a bit and the character of the sound was tweaked in order to make it less annoying.

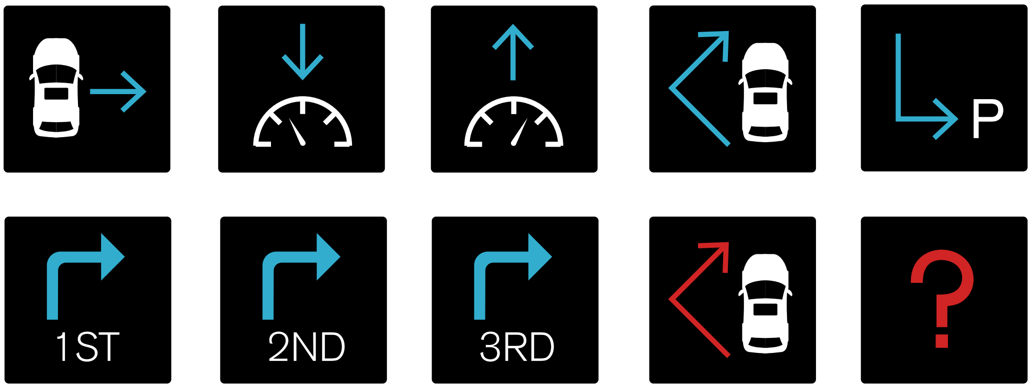


Figure 23. The symbols for feedback. The symbol for unknown gesture (to the lower right) is accompanied by a text encouraging the user to try again.

To sum up, the system we have designed enables the user to perform the gestures presented in Fig. 24 which will influence the driving. When a gesture is performed, immediate multimodal feedback will be communicated to the user regarding if the gesture is recognized or not, and if it is recognized the feedback will convey which instruction that was given. If the gesture is recognized by the system, the confirming sound will be played and the visual feedback will be presented in the DIM (see Fig. 26). In the case of second and third exit/turn, the system will automatically count down as the vehicle passes one exit. However, there might be so that an instruction is not executable, for instance if the user initiates an overtake and there no vehicle to overtake. Then the auditory feedback will be the same as for the unrecognizable gesture, but the visual feedback will be different (see Fig. 26). Besides the fact that the arrow is red instead of blue, there will also be a text message explaining why the overtake is not possible.

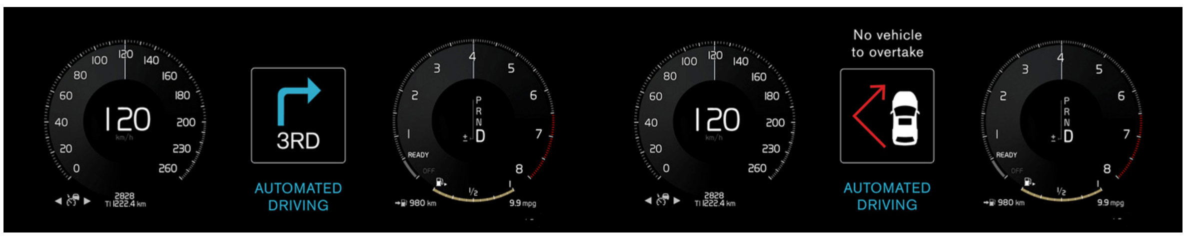


Figure 24. Two examples of the visual feedback in the DIM. The left one shows the feedback for the instruction "take third exit/turn". The right one depicts the feedback that is displayed when the user have given the instruction "overtake" and there is no vehicle to overtake.

7. Discussion

In this chapter we will discuss our study, first in relation to the process then to the final design. Lastly, we discuss and provide suggestions for future work.

7.1. Process

In the beginning of this study we found it a bit difficult to grasp the subject of gestures and how to design them, since we had no previous experience of it. Furthermore, as the area of gestural interaction in autonomous vehicles is rather unexplored with only a few studies in related domains (e.g. Niemann, Peterman and Manzey, 2011; Hinrichs and Carpendale, 2011; Morris, Wobbrock and Wilson, 2010; Lundgren and Hjulström, 2011) we assumed that the design work would be explorative and challenging. It was important to do a fairly extensive literature review on gestures and autonomous driving in order to understand the domain we were expected to design for. The literature provided knowledge and insights which we could apply in the design process. The most important insight we got was the lack of standards and conventions for gestural interfaces (Norman and Nielsen 2010; Norman 2010; Norman and Wadia 2013), which made us realize the importance of involving users throughout the process. This has also been stated by Andersson and Wikander (2013) who suggest that one should ask users before designing the gestures. We would argue that it was essential for us to involve as many as 15 users in the gesture design sessions since we needed to collect a rich set of data in order to explore patterns for which gestures the users found most natural. Without any patterns, how would we know which gesture would be the most natural for the users? It would have been a difficult question to answer.

Continuing to the methods we used in the first iteration, we believe that the order of each use case in the first gesture design session, may have influenced the result. Since we began with the use cases that can be regarded as the most simple and straightforward ones (change lane and increase/decrease speed), the participants might have performed the most natural gestures for these use cases, leaving out the possibility to perform them for other use cases later on in the scenario. However, it could also have been so that by starting with the least complex use cases and then continuing with more complex ones in the scenario, the participants might have performed the most natural gesture every time. Further on the first gesture design session, we believe that the validity of participants mapping of predefined gestures to the different use cases could be questioned since it was performed outside of a vehicle context. Moreover, the gestures were predefined and we did not ask the users whether they found the predefined gestures or the ones they did themselves most natural. Therefore, we regarded the user designed gestures as more important than the data from the gesture mapping.

The patterns from the gesture design session was the basis for the two concepts we designed in the second iteration. However, here we faced the problem of having several patterns that was equally common among the users. To solve this we could have seeked explicit input from users, but we decided to select the gestures we found most natural and simple to perform while at the same time considering the pattern across the gestures, which is in line with the guideline *Use patterns* by Andersson and Wikander (2013). We would argue this was a design decision we had to make in order to progress in our work. Further into the second iteration, we also gained an insight regarding the material aspects of the input device. The friction of the surface of the input device seem to affect the size of the gestures that the user performs. A gesture that is performed on a surface with little friction is, based on our experience, usually of smaller size than a gesture performed on a surface with more friction. However, since we have not conducted any studies of this we do not know whether it is generalizable or not. Still, we would argue that this finding is important to highlight for further studies.

Proceeding to the third iteration, in which we started with designing feedback in two modalities; visual and auditory. Haptic feedback could have been included as well but due to the limited time for this study we decided not to include it. Furthermore, we would argue that visual and auditory feedback are more common in a vehicle context compared to the haptic modality and that the design space for haptics is fairly complex. Therefore, we designed a simulator with visual and auditory feedback that was contextually aligned with the environment in a physical car. Further, we would argue that by using pre-recorded video and a physical car the test environment became fairly close to actually traveling in a semi-autonomous car. We wanted to achieve this since the context could influence the design of the gestures performed by the users (Hinrichs and Carpendale 2011).

To use pre-recorded video in simulator tests for semi-autonomous driving is something we would argue is favourable since it provides more realism than a computer generated environment and the steering wheel is not going to be used. Additionally, with pre-recorded video the driving style will be the same for every test and the possible differences that can occur by having a *"Wizard of Oz"* driver (Maguire 2001), such as driver fatigue, mistakes or difference between different wizards is eliminated. Yet, the potential drawbacks of this method is that the steering wheel in the physical vehicle was static and the video recording was divided into short clips rather than one single continuous driving sequence. However, we would argue that these factors did not influence the result extensively.

7.2. Final Design

By introducing semi-autonomous driving the user's mental model for the task of driving will break, since instead of using traditional hardware in terms of pedals and a steering wheel they will be obligated to interact in a new manner while the vehicle drive autonomously. The previous natural interaction with hardware is changed to performing multi-touch gestures on an input device which cause a disruption in the user'ss mental model. A similar disruption occurred when touch screens were first introduced (Bradley 2011), since it forced the users to interact in new ways and form new mental models for how such system functions. In this study we have searched for the most natural interactions and incorporated these in the conceptual model in order to facilitate the user's process of forming a mental model of the system.

The most natural gestures seem to be those that are analog (Zhai et al. 2011). All of the gestures we have designed are analog, except the gesture for undo which is more abstract. However, since the gesture for undo does not affect the driving but rather a pending instruction we would argue that it is suitable for it to be abstract. A double tap gesture instead of a swiping gesture, makes it more differentiated from the other gestures and easier to perform which is in line with the guideline Easy undo by Andersson and Wikander (2013). Yet, the double-tap is often used for selecting an object in other contexts which might be confusing. Moreover, the gestures we have designed are all single finger gestures, except for the use cases take second/third exit/turn, which is in line with the findings of the study by Morris, Wobbrock and Wilson (2010). Furthermore, the gestures for take second/third exit/turn are variations of the gesture for take the first exit/turn which is inline with the guideline Use patterns by Andersson and Wikander (2013).

In the search for natural interactions we also considered the importance of affordances and since the input device does not make visible its affordance of recognizing gestures, we decided to explore the possibilities of having gesture hinting (Lundgren and Hjulström 2011) as signifiers (Norman 2013). Our assumption was that it would help the user to discover possible interactions, which otherwise is difficult in gesture-based interfaces (Norman 2010). However, since the result from our evaluation indicated that the users seem to disregard the hints and had trouble understanding them, we decided not to include it in our design. By doing that the system still face the problem of a lack of perceivable affordances (Saffer 2008; Norman 2010; Norman and Nielsen 2010). Yet, even if it might be problematic we would argue that after the the initial learning phase the users will know the possible interactions and the need of affordances is reduced. This would be similar to the findings of Butz, Kramer and Rümmelin's study in 2013 in that gestural interfaces outperforms button interfaces (which often have more apparent affordances) after the learning phase.

Continuing to the feedback, a positive aspect in our system is that it provides immediate feedback. This has been presented as essential for gestural interfaces (Norman and Wadia 2013) and is also defined by Saffer (2008) as the characteristic *High responsivity*. Furthermore, our findings indicate that the multimodality facilitate the user's understanding of the system which is in line with the research by Pfleging et al. (2011) who argue that by employing multimodality the advantages can outweigh the disadvantages of every modality.

Lastly, by introducing semi-autonomous driving and the system we have designed, we would argue that the role of the driver is shifting from being an operator towards becoming a supervisor similar to the findings of the study by Damböck et al. (2012). It means that the potential out-of-the-loop problems (Endsley and Kiris 1995; Niemann, Peterman, and Manzey 2011) could be prominent and it is something that we have not specifically studied. However, we would argue that since the system provides a possibility for the driver to influence the task of driving, it also provides an incentive for the driver to be in-the-loop and thereby decrease the risk of out-of-the-loop problems.

7.3. Future Work

The system we have designed in this study is one of the first steps towards supporting semi-autonomous driving with gesture-based interfaces. However, there are several aspects that needs to be further studied. The most prominent one would be to study if there is a need to queue instructions. For instance, after giving the instruction to take the next exit/turn, the user might want to give the vehicle the instruction to take the next exit/turn to the left. This emerged during the third iteration, but since we could not observe users driving freely with our system it was difficult to find any basis for the need of a queue. However, we would argue that queueing instructions might very well be needed. Therefore we propose a suggestion for a queue which makes this possible (see Fig. 25). The design is based on the metaphor of seeing road signs along the road, which is something drivers are used to do. The pending instruction is placed closer to the driver and the upcoming instructions are further away.

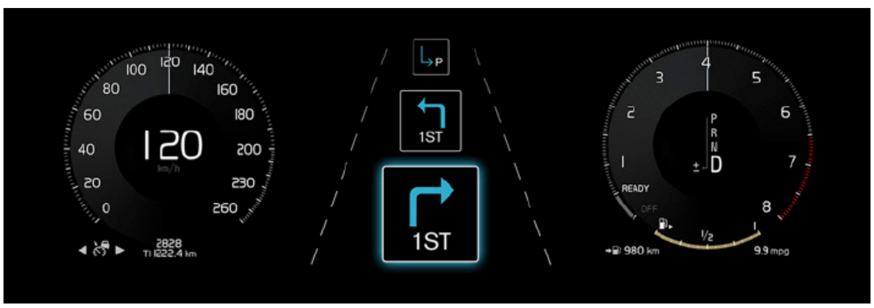


Figure 25. The proposed design for the queue of instructions. The lowest and largest instruction is the first to be executed whilst the instruction in the top is the last.

Regarding the placement of the visual feedback several participants in our final user test expressed a desire for it to be placed in the HUD instead of the DIM, since they wanted to be able to keep their eyes on the road. However, according to Hancock et al. (1999), in the future drivers will be more engaged in other activities (e.g. calling and texting) while driving and we believe it will increase even more with autonomous vehicles. Furthermore, as the vehicle drive autonomously there will be less need for drivers to keep their eyes on the road. We would argue that by placing the feedback in the HUD (Head-Up Display) the drivers are forced to look out the window to view the feedback and indirectly perceive the surrounding environment. That could increase situational awareness and decrease out-of-the-loop problems. Yet, it needs to be further studied in order to find out if there is a real need for placing the feedback in the HUD.

Additionally, we believe it would be advantageously to merge the vehicle's navigation system with our system. That could increase the driver's level of involvement in the task of driving since the navigational system could display the next action by the vehicle even if no instruction is pending. Furthermore, when the driver gives the vehicle an instruction the change of route could be displayed. It could also provide additional information of the current route and traffic situation. Another benefit of this merge could be that the navigation system could provide further means of entering destinations for the user. Destinations could be entered by gesturing letters on the input device. For instance, by performing a gesture shaped like a "G" would be equal to pressing the G-button on a keyboard.

Continuing to the use cases employed in our system, we have realized that the use cases for increase and decrease speed might not be what the driver of the future really needs. Since autonomous vehicles most likely will keep the speed limit, the use of changing speed might be redundant. Furthermore, in reality it could be more plausible that the goal of the driver is to make sure that the destination is reached on time, or that the ride is as fuel efficient as possible. Therefore we would argue that there is a need to investigate these use cases more thoroughly to get a better understanding of the driver's goal with increasing or decreasing the speed. For instance, the uses cases could be altered to control the driving style of the vehicle rather than the speed.

Regarding the gestures we had an idea that the system could support variations of gestures for the same use case, similar to the findings of Hinrichs and Carpendale's study in 2011. That would make the system more forgiving and support the preferences of different users. Furthermore, the system could allow customization such as letting the users define their own gestures for each use case if they find a particular gesture to be unnatural or badly designed. That is also suggested by Andersson and Wikander in the guideline *Allow customization*. Additionally, increased support for multi-touch could allow the user to evolve into an expert user similar to the guideline *Allow the user to evolve into an advanced user* by Andersson and Wikander (2013). For instance, performing the change lane or overtake gesture with two fingers could mean that the vehicle should change two lanes or overtake two vehicles.

Since there was a quite large spread among the participants' preferences regarding the position of the input device, the placement could be further studied. Even though our findings pointed towards placing the input device between the armrest and the gear stick, we suggest that it could be more seamlessly integrated into the vehicle. There could be multiple input areas where the gestures can be performed, such as in the steering wheel or on the interior of the door next to the driver. However, if this would to be applied, it is important that the input areas communicate that they are interactive. Otherwise it could be difficult for the users to understand where they can interact, which has been previously observed in the study by Fang and Ainsworth (2012). Furthermore, it would be important to keep the trace of the gesture which we had in our prototype.

8. Conclusion

In this study we have designed a system for semi-autonomous driving which includes a set of gestures performed on an input device, supported by visual and auditory feedback. We have addressed the research question of which properties a multi-touch gestural interaction system for semi-autonomous driving should have. Furthermore, our design process followed a user-centric design approach throughout the study which resulted in a design based on user data from several iterations. This has been particularly important since the field of gestural interaction in autonomous vehicle is largely unexplored.

To conclude, this study suggests that the conceptual model for gesture-based semi-autonomous driving should be founded in analog gestures (Zhai et al. 2011) aiming to resemble the movement of the vehicle. Furthermore, our findings indicate that the gestures should be designed to be as simple as possible in terms of gestural motion, even if a more complex gesture could be more similar to the movement of the vehicle. For instance, a swipe to the right seems to be easier to perform than a swipe upwards and to the right. The gestures should also be continuous and there should be a consistent pattern across the entire set of gestures, which has been previously stated by Andersson and Wikander (2013). Additionally, our findings point towards that single finger gestures are easier to perform compared to multi-touch gestures, similar to the findings of Morris, Wobbrock and Wilson's (2010) study. Lastly, we conclude that it is essential that the system is forgiving. If the user would perform a recognizable gesture slovenly, it should still be confirmed by the system.

Further, our findings suggests that instantaneous feedback is important, which has been previously stated by Norman and Wadia (2013), in order for the users to understand what happens when an instruction is initiated. Moreover, the multimodality with both visual and auditory feedback seem to be more efficient and helpful compared to using only one modality, which is in line with the findings from the study by Pfleging et al. (2011). It could be due to that the auditory feedback communicates a response to the user's action whilst the visual feedback communicates the details of the response. Together they convey a richer message to the user than if the specific modalities would have been used separately. Lastly, we conclude that the feedback should be connected to the shape of the gesture as that could facilitate learning and remembering the gestures.

Our work has contributed several findings to the area of gestural interaction for semi-autonomous driving. However, more studies has to be conducted in order to explore this field more extensively and investigate the future work we have proposed. Moreover, the system needs to be studied over a longer period of time, i.e. beyond the learning phase, in order to better understand its potential and possible issues. This study is just one of the first steps towards introducing a new way of driving, with gestural input for initiating instructions to a semi-autonomous vehicle.

References

Abrahamson, H. (2014). Googles Robotbil Klar för Provkörning. Available at: http://www.nyteknik.se/nyheter/fordon_motor/bilar/article3873486.ece. Accessed at: 2015-02-02

Andersson, D. and Wikander, F. (2013). Designing In-vehicle Gestural Interfaces - Minimising visual distraction. Master Thesis. Chalmers Tekniska Högskola.

Barker, A. (1997). 30 minutes...to Brainstorm Great Ideas. [Books24x7 version] Available at: http://common.books24x7.com.proxy.lib.chalmers.se/toc.aspx?bookid=6546. Accessed at: 2015-02-02

Benyon, D. (2010). Envisonment. Chapter 8. In: Designing Interactive Systems. pp 177-197. Addison-Wesley.

Beyer, H., and Holtzblatt, K. (1998). Contextual Design: Defining Customer-centered Systems. San Francisco, Calif.: Morgan Kaufmann. pp. 154-163.

Billings, C. E. (1996). Human-centered Aviation Automation: Principles and Guidelines.

Billings, C. E. (1997). Aviation Automation: The Search For a Human-centered Approach.

Bradley, S. (2011). 3 Representational Models That Affect Usability. Available at: http://www.vanseodesign.com/web-design/representational-models/. Accessed at: 2015-02-02

Brown, T. (2008). Design Thinking. Harvard Business Review, 86(6), pp. 84-92.

Butz, A., Kroner, V. and Rümmelin, S. (2013). Simple Nonvisual Interaction on Touch Tablets. In: Proceedings of Mensch und Computer 2013, Sept 8 - 11, 2013, Bremen, Germany. pp. 241-250.

Chalmers (2015). Available at: http://www.chalmers.se/safer/EN/projects/pre-crash-safety/associated-projects/aimmit-automotive. Accessed at: 2015-01-27.

Cooper, A., Reimann, R., Cronin, D., and Noessel, C. (2014). About Face.

Damböck, D., Weißgerber, T., Kienle, M., & Bengler, K. (2012). Evaluation of a Contact Analog Head-Up Display for Highly Automated Driving. In: Advances in Human Aspects of Road and Rail Transportation. Stanton, N. A. (Ed.). (2012).

Encyclopedia Britannica, (2015). Mapping. Available at: http://global.britannica.com/EBchecked/topic/363594/mapping. Accessed at: 2015-03-17

Endsley, M. R. (1988). Situation Awareness Global Assessment Technique (SAGAT). Aerospace and Electronics Conference, 1988. NAECON 1988., Proceedings of the IEEE 1988 National. IEEE, 1988.

Endsley, M., and Kiris, E. (1995). The Out-of-the-Loop Performance Problem and Level of Control in Automation. In: Human Factors, 37(2). pp. 381-394. doi:10.1518/001872095779064555

Fang, V., and Ainsworth, L. (2012) Designing & Rapid Prototyping a Gesture-Enabled Steering Wheel. In: Adjunct Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI'12), October 17-19, 2012, Portsmouth, NH, USA. pp. 53-56.

Fullerton, T. (2014). Game Design Workshop. Boca Raton: Taylor & Francis.

Gibson, J.J. The Theory of Affordances. Hilldale, USA (1977).

Gold, C., Dambock, D., Lorenz, L., and Bengler, K. (2013). “Take over!” How long does it take to get the driver back into the loop?. In: Proceedings Of The Human Factors And Ergonomics Society Annual Meeting, 57(1). pp 1938-1942. doi:10.1177/1541931213571433

Google Self-Driving Car Project. (2014). Google Self-Driving Car Project. Available at: https://plus.google.com/+SelfDrivingCar/posts/9WBWP2E4GDu. Accessed at: 2015-02-02

Gould, J. D. (1988). How to Design Usable Systems. In: Handbook of Human-Computer Interaction. (Helander, M.). pp. 757-789.

Greca, I. and Moreira, M. (2000). Mental Models, Conceptual Models, and Modelling. In: International Journal of Science Education. 22(1). pp. 1-11.

Gulliksen, J. Lanz, A. Boivie, I. (1999) User Centered Design - Problems and Possibilities. In: SIGCHI Bulletin. 31(2). pp. 25-35.

Göhring, D., Latotzky, D., Wang, M. and Rojas, R. (2013). Semi-autonomous Car Control Using Brain Computer Interfaces. In: Intelligent Autonomous Systems 12, 194, pp. 393-408.

Hancock, P., Simmons, L., Hashemi, L., Howarth, H., and Ranney, T. (1999). The Effects of In-Vehicle Distraction on Driver Response During a Crucial Driving Maneuver. In: Transportation Human Factors, 1(4) pp. 295-309. doi:10.1207/sthf0104_1

Hinrichs, U., and Carpendale, S. (2011, May). Gestures in the Wild: Studying Multi-touch Gesture Sequences on Interactive Tabletop Exhibits. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 3023-3032. ACM.

Johnson, J. and Henderson, A. (2002). Conceptual Models: Begin by Designing What To Design. In: Interactions Jan/Feb, pp. 25-32.

Lorenz, L., Kerschbaum, P., and Schumann, J. (2014). Designing Take over Scenarios for Automated Driving: How Does Augmented Reality Support the Driver to Get Back into the Loop?. In: Proceedings Of The Human Factors And Ergonomics Society Annual Meeting, 58(1). pp. 1681-1685. doi:10.1177/1541931214581351

Lowdermilk, T. (2013). User-centered Design. Sebastopol, CA: O’Reilly.

Lundgren, S., and Hjulström, M. (2011, September). Alchemy: Dynamic Gesture Hinting for Mobile Devices. In: Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments. pp. 53-60. ACM.

Maguire, M. (2001). Methods to Support Human-centred Design. In: International Journal Of Human-Computer Studies, 55(4). pp. 587-634. doi:10.1006/ijhc.2001.0503

Morris, M. R., Wobbrock, J. O., and Wilson, A. D. (2009, April). User-defined Gestures for Surface Computing. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 1083-1092. ACM.

Morris, M. R., Wobbrock, J. O., and Wilson, A. D. (2010, May). Understanding Users' Preferences For Surface Gestures. In: Proceedings of graphics interface 2010. pp. 261-268. Canadian Information Processing Society.

Niemann, Petermann and Manzey. (2011). Driver In The Loop: Manoeuvre Based Driving as an Automation Approach. In: Human Centred Automation. (Gérard, N., Manzey, D., Onnasch, L., Waard, D. and Wiczorek, R.)

Nielsen, M., Störring, M., Moeslund, T. B., and Granum, E. (2003, March). A Procedure for Developing Intuitive and Ergonomic Gesture Interfaces for HCI. In: Proceedings of the 5th International Gesture Workshop. pp. 1-12.

Norman, D. A. (1999). Affordance, Conventions, and Design. In: Magazine interactions, 6(3), pp. 38-43.

Norman, D. A. and Nielsen (2010). Gestural Interfaces: A Step Backward In Usability.

Norman, D. A. (2006). Logic Versus Usage: The Case for Activity-centered Design. In: Magazine interactions, 13(6), pp. 45, 63.

Norman, D. A. (2010). Natural Interfaces Are Not Natural. Available at: http://www.jnd.org/dn.mss/natural_user_interfa.html. Accessed at 2015-01-21.

Norman, D. A. and Wadia, B. (2013). Opportunities and Challenges For Touch and Gesture-Based Systems. Available at: http://www.jnd.org/dn.mss/opportunities_and_ch.html. Accessed at 2015-01-21.

Norman, D. A. (1983). Some Observations on Mental Models. Edited by Genter and Stevens. Published by Psychology press (2014).

Norman, D. A. (2013). The Design of Everyday Things, Revised and Expanded Edition. [Books24x7 version] Available at: http://common.books24x7.com.proxy.lib.chalmers.se/toc.aspx?bookid=59487. Accessed 2015-02-02.

Normark, C. J. (2014) The User as Interface Designer, Ph.D-thesis, Luleå Technical University.

Pfleging, B., Kienast, M., Schmidt, A., and Döring, T. (2011). SpeeT: A Multimodal Interaction Style Combining Speech and Touch Interaction in Automotive Environments. In: Adjunct Proceedings of the 3rd International Conference on Automotive User Interfaces and Interactive Vehicular Applications, AutomotiveUI, 11.

Rahman, A., Saboune, J. and El Saddik, A. (2011). Motion-path Based in Car Gesture Control of the Multimedia Devices . In: Proceedings of the First Acm International Symposium on Design and Analysis of Intelligent Vehicular Networks and Applications - DIVANet '11.

Redström, J. (2006) Towards User Design? On the Shift From Object to User As the Subject of Design. In: Design Studies, 27(2) (2006) pp. 123-139.

Rosson, M., B. and Carroll, J., M.. (2002). Usability Engineering: Scenario-based Development of Human-computer Interaction. [Books24x7 version] Available at: http://common.books24x7.com.proxy.lib.chalmers.se/toc.aspx?bookid=3150

Saffer, D. (2008). Designing Gestural Interfaces. Sebastopol, Calif.: O'Reilly Media.

Sauer, J., and Sonderegger, A. (2009). The Influence of Prototype Fidelity and Aesthetics of Design in Usability Tests: Effects on user behaviour, subjective evaluation and emotion. In: Applied Ergonomics, 40(4), pp. 670-677. doi:10.1016/j.apergo.2008.06.006

Scheiben, A., Temme, G., Köster, F. and Flemisch, F. (2011). How To Interact With a Highly Automated Vehicle - Generic Interaction Design Schemes and Test Results of a Usability Assessment. In: Human Centred Automation. (Gérard, N., Manzey, D., Onnasch, L., Waard, D. and Wiczorek, R.)

Semcon (2015). Detta är Semcon. Available at: http://www.semcon.com/sv/Om-Semcon/ Accessed at: 2015-06-15.

Sharp, H., Rogers, Y. and Preece, J. (2007). Interaction Design: Beyond Human-Computer Interaction. Chichester: Wiley & Sons.

Taylor, E. and Oreskovic, A. (2015). Apple Studies Self-driving Car, Auto Industry Source Says. Reuters. Available at: http://www.reuters.com/article/2015/02/14/us-apple-autos-idUSKBN0LI0IJ20150214. Accessed at: 2015-02-16.

Toolkit, HCD. 2nd edition. (2008) IDEO.

Virzi, R. (1989). What Can You Learn From a Low-Fidelity Prototype? In: Proceedings Of The Human Factors And Ergonomics Society Annual Meeting, 33(4), pp. 224-228. doi:10.1177/154193128903300405

Volvo Cars. (2013). Volvo Car Group Initiates World Unique Swedish Pilot Project With Self-driving Cars on Public Roads. Available at: https://www.media.volvocars.com/global/en-gb/media/pressreleases/136182/volvo-car-group-initiates-world-unique-swedish-pilot-project-with-self-driving-cars-on-public-roads. Accessed at: 2015-02-02.

Williams, A. (2009). User-centered Design, Activity-centered Design, and Goal-directed Design: A Review of Three Methods for Designing Web Applications. In: Proceedings of the 27th ACM international conference on Design of communication (SIGDOC '09), pp. 1-8. ACM, New York, NY, USA.

Wilson, C (2013), Chapter 1 - Brainstorming, In: Brainstorming and Beyond, edited by Chauncey Wilson, Morgan Kaufmann, Boston, 2013, pp. 1-41, ISBN 9780124071575. Available at: http://dx.doi.org/10.1016/B978-0-12-407157-5.00001-4.

Yamauchi, Y. (2012). Participatory Design. In: Field Informatics: Kyoto University Field Informatics Research Group. Springer Science & Business Media. Ishida, T. (Ed.). (2012)

Zhai, S., Kristensson, P. O., Appert, C., Anderson, T. H. and Cao, X. (2011). Foundational Issues in Touch-Surface Stroke Gesture Design - An Integrative Review. In: Foundations and Trends in Human Computer Interaction, 5(2). pp. 97-205.

Zhao, T. and McDonald, S. (2010). Keep Talking. In: Proceedings of the 6th Nordic Conference on Human-Computer Interaction Extending Boundaries - NordiCHI '10.

Appendices

A - Scenario for the first gesture design session

1. You are on your way to work, driving in the right lane on a highway with two lanes. You are approaching a truck which is driving slower that you and want to switch lane to the left one.
2. When you have passed the lorry you want to increase the speed a bit.
3. Then, after a while of driving, the road becomes more winding and you want to decrease the speed to make the driving more comfortable.
4. Thereafter, you want to take the next exit to the right.
5. Then you want to take the second next exit to the left.
6. After you've done that you realize that it's not the second exit you should take, it's the third. Therefore, you need to undo the command. (Ask the participants if they want to have an undo command or just override with a new command).
7. Then tell the car that you want to take the third exit to the left.
8. After the turn you come to a highroad and after a while you approach a car which is driving slower than you. Therefore, you want to overtake it.
9. Then another car is approaching you from behind and you want to let it overtake you. Therefore you want to move a bit to the right inside the lane.
10. After you been overtaken, you realize that you need to make a stop at next roadside/parking place along the road.

B - Scenario for the second gesture design session

1. You are driving on a highway and you want to take the next exit.
2. Having taken the exit you end up behind someone that is practicing driving. Therefore you wish to slow down and keep some extra distance between you and that car.
3. Then someone is approaching you from behind, wishing to overtake you. Therefore you want to move a little bit to the right in your lane to facilitate the overtake.
4. When you have been overtaken you wish to take the second turn ahead.
5. Then you come to a highway, driving in the right lane. After a while you pass a freeway ramp at the same time as someone else entering the highway. In order to be nice and leave space for that driver you want to change to the left lane.
6. When you have switched lane you want to increase speed a bit.
7. Then you want to take the third exit ahead.
8. After the turn you come to a highroad and after a while you approach a car which is driving slower than you. Therefore, you want to overtake it.
9. After a while you want to stop at the next parking place along the road for some fresh air.
10. But then you realize that it could wait so you want to undo the last instruction.

C - Questionnaire for final evaluation

Ålder:____________

Man:___   Kvinna:___

År med körkort:__________

Hur ofta kör du bil?
[ ] Varje dag   [ ] Flera ggr/v   [ ] 1-2 ggr/v   [ ] Nån gång/mån   [ ] Nån gång/år
[ ] Aldrig

Frågor

- Hur upplevde du att använda systemet?
- Vad tyckte du var bra/dåligt med systemet?
- Hur svårt var systemet att använda på en skala från 1 till 10?
  Varför?
- Upplevde du några problem med systemet när du använde det?
  Beskriv problemet.
  Varför?
- Vad tyckte du om gesterna?
  Motivera gärna.
- Kändes någon/några gester onaturliga?
  Varför?
- Var gesterna svåra att komma ihåg?
  Varför?
- Vad tyckte du om antalet gester?
  Varför?
- Vad tyckte du om att använda ett finger för alla gester förutom andra/tredje avfarten?
  Varför?
- Vad tyckte du om att använda två/tre fingrar för andra/tredje avfarten?
  Varför?
- Vad tyckte du om att byt fil / flytta sig åt sidan var samma gest?
  Varför?
- Vad tyckte du om ångra gesten?
  Varför?
- Vad tycker du om att ångra gesten skiljer sig mot de andra gesterna?
- Hur visste du att du gjort rätt gest?
- Vilken typ av feedback fick du?
  Var fanns feedbacken?
- Använde du dig av feedbacken på gesttableten?
  Varför?
- På vilket sätt?
  Uppfattade du den pulserande ramen runt feedbacken?
  Förstod du den?
- Förstod du symbolerna för feedback?
  Varför?
- Var det någon/några symboler som var svåra att förstå?
  Varför?
- Hur upplevde du ljuden för feedback?
  Varför?

- Var det något ljud otydligt?
- Saknade du någon feedback?
  Varför?
- Skulle du vilja ändra på något?