Realistic Interaction with Virtual Objects Within Arm's Reach

by Mathias Moehring

A dissertation submitted in confirmity with
the requirements for the degree of Dr.-Ing.

Wolfsburg, October 2012

Abstract

The automotive industry requires realistic virtual reality applications more than other domains to increase the efficiency of product development. Currently, the visual quality of virtual environments resembles reality, but interaction within these environments is usually far from what is known in everyday life. Several realistic research approaches exist, however they are still not all-encompassing enough to be usable in industrial processes. This thesis realizes lifelike direct multi-hand and multi-finger interaction with arbitrary objects, and proposes algorithmic and technical improvements that also approach lifelike usability. In addition, the thesis proposes methods to measure the effectiveness and usability of such interaction techniques as well as discusses different types of grasping feedback that support the user during interaction.

Realistic and reliable interaction is reached through the combination of robust grasping heuristics and plausible pseudophysical object reactions. The easy-to-compute grasping rules use the objects' surface normals, and mimic human grasping behavior. The novel concept of Normal Proxies increases grasping stability and diminishes challenges induced by adverse normals. The intricate act of picking-up thin and tiny objects remains challenging for some users. These cases are further supported by the consideration of finger pinches, which are measured with a specialized finger tracking device. With regard to typical object constraints, realistic object motion is geometrically calculated as a plausible reaction on user input. The resulting direct finger-based interaction technique enables realistic and intuitive manipulation of arbitrary objects.

The thesis proposes two methods that prove and compare effectiveness and usability. An expert review indicates that experienced users quickly familiarize themselves with the technique. A quantitative and qualitative user study shows that direct finger-based interaction is preferred over indirect interaction in the context of functional car assessments. While controller-based interaction is more robust, the direct finger-based interaction provides greater realism, and becomes nearly as reliable when the pinch-sensitive mechanism is used.

At present, the haptic channel is not used in industrial virtual reality applications. That is why it can be used for grasping feedback which improves the users' understanding of the grasping situation. This thesis realizes a novel pressure-based tactile feedback at the fingertips. As an alternative, vibro-tactile feedback at the same location is realized as well as visual feedback by the coloring of grasp-involved finger segments. The feedback approaches are also compared within the user study, which reveals that grasping feedback is a requirement to judge grasp status and that tactile feedback improves interaction independent of the used display system. The considerably stronger vibrational tactile feedback can quickly become annoying during interaction.

The interaction improvements and hardware enhancements make it possible to interact with virtual objects in a realistic and reliable manner. By addressing realism and reliability, this thesis paves the way for the virtual evaluation of human-object interaction, which is necessary for a broader application of virtual environments in the automotive industry and other domains.

Zusammenfassung

Stärker als andere Branchen benötigt die Automobilindustrie realistische Virtual Reality Anwendungen für eine effiziente Produktentwicklung. Während sich die visuelle Qualität virtueller Darstellungen bereits der Realität angenähert hat, ist die Interaktion mit virtuellen Umgebungen noch weit vom täglichen Erleben der Menschen entfernt. Einige Forschungsansätze haben sich mit realistischer Interaktion befasst, gehen aber nicht weit genug, um in industriellen Prozessen eingesetzt zu werden. Diese Arbeit realisiert eine lebensnahe mehrhändige und fingerbasierte Interaktion mit beliebigen Objekten. Dabei ermöglichen algorithmische und technische Verbesserungen eine realitätsnahe Usability. Außerdem werden Methoden für die Evaluation dieser Interaktionstechnik vorgestellt und benutzerunterstützende Greiffeedbackarten diskutiert.

Die verlässliche und gleichzeitig realistische Interaktion wird durch die Kombination von robusten Greifheuristiken und pseudophysikalischen Objektreaktionen erreicht. Die das menschliche Greifverhalten nachbildenden Greifregeln basieren auf den Oberflächennormalen der Objekte. Die Reduktion negativer Einflüsse verfälschter Normalen und eine höhere Griffstabilität werden durch das neuartige Konzept der Normal Proxies erreicht. Dennoch bleibt für manche Nutzer das Aufnehmen von dünnen und kleinen Objekten problematisch. Diese Fälle werden zusätzlich durch die Einbeziehung von Fingerberührungen unterstützt, die mit einem speziellen Fingertracking Gerät erfasst werden. Plausible Objektreaktionen auf Benutzereingaben werden unter Berücksichtigung typischer Objekteinschränkungen geometrisch berechnet.

Die Arbeit schlägt zwei Methoden zur Evaluierung der fingerbasierten Interaktion vor. Ein Expertenreview zeigt, dass sich erfahrene Benutzer sehr schnell in die Technik einfinden. In einer Benutzerstudie wird nachgewiesen, dass fingerbasierte Interaktion im hier untersuchten Kontext vor indirekter Interaktion mit einem Eingabegerät bevorzugt wird. Während letztere robuster zu handhaben ist, stellt die fingerbasierte Interaktion einen deutlich höheren Realismus bereit und erreicht mit den vorgeschlagenen Verbesserungen eine vergleichbare Verlässlichkeit.

Um Greifsituationen transparent zu gestalten, realisiert diese Arbeit ein neuartiges druckbasiertes taktiles Feedback an den Fingerspitzen. Alternativ wird ein vibrotaktiles Feedback am gleichen Ort realisiert und visuelles Feedback durch die Einfärbung der griffbeteiligten Fingersegmente umgesetzt. Die verschiedenen Feedbackansätze werden in der Benutzerstudie verglichen. Dabei wird Greiffeedback als Voraussetzung identifiziert, um den Greifzustand zu beurteilen. Taktiles Feedback verbessert dabei die Interaktion unabhängig vom eingesetzten Display. Das merklich stärkere Vibrationsfeedback kann während der Interaktion störend wirken.

Die vorgestellten Interaktionsverbesserungen und Hardwareerweiterungen ermöglichen es, mit virtuellen Objekten auf realistische und zuverlässige Art zu interagieren. Indem die Arbeit Realismus und Verlässlichkeit gleichzeitig adressiert, bereitet sie den Boden für die virtuelle Untersuchung von Mensch-Objekt Interaktionen und ermöglicht so einen breiteren Einsatz virtueller Techniken in der Automobilindustrie und in anderen Bereichen.

Acknowledgements

There are many outstanding individuals who contributed greatly to the success of this thesis, and I would like to take the opportunity to express my utmost gratitude for their unwaivering support.

First and foremost, I want to thank Bernd Froehlich who served as my mentor throughout this entire thesis. His outstanding support and direction helped guide me to a consequent evolution of the direct finger-based interaction metaphor, and always pushed the limits of the work so it could reach its full potential. I also want to thank Gabriel Zachmann who unhesitatingly agreed to review this thesis.

Peter Zimmermann, the former head of the Volkswagen VR research group, contributed a countless number of ideas in the early stages that served as the inspirational basis of this work and particularly helped facilitate the development of the tactile grasping feedback. I owe a special thanks to his successor Ralf Rabaetje, who constantly supported my thesis with motivating and constructive criticism. I also want to thank Werner Schreiber who strongly encouraged and motivated me to preserve and finish this work.

I owe a multitude of thanks to the Volkswagen research team, who always provided an incomparable working atmosphere and was always available for an endless number of technical discussions. Their tireless willingness to evaluate the developed technologies was paramount. From the team, I want to especially mention Ralph Voges, Jan Jacobs and Eduard Jundt for their support. Sincere thanks are also extended to Konrad Zuerl and the Advanced Realtime Tracking team for their creative and reliable realization of the hardware ideas. I also want to thank Frank Ostermann and Martin Bergmann for their support and motivation thorughout the final phase of the thesis.

The work could not have been completed without the students that supported the development and evaluation of the techniques. Special thanks to Robert Scheibe who made the realization of the tactile feedback system possible with his enormous technical know-how.

I owe a debt of gratitude to my parents who made this work possible because of their love, support and patience. And finally, I want to thank Susanne who served as my rock of support and well of understanding in agreeing to invest all the time necessary to make this work a success. Thank you.

Contents

Chapter 1

Introduction

1.1 Motivation

The automotive industry continuously participates in Virtual Reality (VR) research to benefit from the potential of this technology. To a large extent, automotive product development processes are characterized by trade-offs induced by diverging customer demands and economic requirements. On one hand, customers increasingly expect highly individualized products made with outstanding quality. This leads to a large variety of complex car configurations that are time-consuming to produce and cost-intensive to develop. On the other hand, intense competition with decreased profit margins and fast moving innovation cycles create cost and time pressure and demand compressed development processes.

Virtual Reality is a technology used to diminish the conflict between profit and demand. The development processes in the automotive industry are increasingly based on digital car data due to the evolution of Computer-Aided Design (CAD). Virtual Reality provides developers the ability to use digital car data throughout the development and assessment processes necessary for product development. Virtual Reality assessments are easily conducted in the very early stages of development before real hardware mockups are built. The ability to virtually evaluate car models significantly reduces the number of needed product prototypes, and helps designers to detect construction challenges early on, which leads to development savings and an increase in product quality. Moreover, digital car data is flexible unlike static hardware mockups. This flexibility allows developers to conduct assessments that involve variants and design alternatives.

Nowadays, high-quality visualizations of virtual car models are commonly used throughout the whole product process (e.g. for applications concerning the comparison of design alternatives, media production, and marketing). In these computer-generated images, only the visual characteristics of a newly developed car are of interest, and the functional aspects are not considered at this time. Few interactive applications that focus on functional characteristics of virtual car models are established. The few applications available do not require a high level of realism. For example, assembly simulation applications implement abstract interaction metaphors based on input devices to evaluate assembly sequences and tool clearances without focusing the interplay between the human being and the car.

However, a high number of assessments in the development processes deal with the human-car interaction and the functional characteristics of the car. For the realization of functional assessments based on virtual car models, it is inevitable to have interactive immersive systems with realistic interaction metaphors that are close to reality. Otherwise, the results of the assessments cannot be transferred to reality. The substitution of a mockup-based assessment by a Virtual Reality application can only succeed if this application is easy to use and works robustly. Consequently, interaction metaphors for functional assessments of virtual car configurations have to be plausible and reliable at the same time.

Plausibility and reliability are the motivations for the development of a variety of interaction techniques, however these requirements often lead to diverging solutions. Robust and efficient object manipulation is often the major goal for research approaches without any further demands on the interaction metaphor. As a result, indirect interaction metaphors were developed. Indirect interaction metaphors incorporate adequately designed input devices or widgets that provide robust and easy-to-learn interaction techniques. These indirect metaphors based on buttons and controllers provide only a limited amount of realism.

In contrast to indirect interaction metaphors, the use of virtual representations of the user's body, hands and fingers, allows for very realistic and direct object manipulations that closely mimic real ones. In particular, the dexterity of fingers and the different ways humans grasp objects play a central role in how drivers interact with a car. Unfortunately, a huge difference with respect to robustness and efficiency between real and virtual finger-based interactions exists due to the limited sensory feedback in virtual environments. Moreover, these techniques come with the same limitations that humans encounter in the real world, such as the limitation of interaction space.

Because of the limitations, interaction metaphors for virtual reality applications are subject to a trade-off. On the one hand, interaction metaphors can be very efficient, all-encompassing and robust. The robustness is often achieved by using indirect approaches employing controller-based input devices. On the other hand, realistic interaction using direct finger-based methods suffers from the limited sensory feedback of virtual reality systems, and thus have traditionally been less robust than indirect methods or their real-world counterparts.

All the observations discussed thus far in this thesis have led to the central research question: how can direct finger-based interaction in virtual environments be made more robust while remaining realistic at the same time? The following considerations are investigated in order to answer this question:

1. What are the algorithmic and technical improvements for direct finger-based interaction that make the usability acceptable for a variety of automotive evaluation tasks?
2. How to measure the effectiveness and usability of direct finger-based interaction in comparison to other techniques?
3. Which types of grasping feedback are helpful and preferred for performing finger-based interaction tasks?

1.2 Applications

This chapter provides an overview of the possible application areas, and the resulting requirements for using direct finger-based interaction in functional validations of virtual car interiors. These validations deal with the ergonomic aspects of a car's interior as well as the usability of human-car interfaces. Typical car attributes that are analyzed with these assessments are listed in table 1.1.

At present, first virtual assessments of functional aspects are carried out on a purely visual basis with static models in immersive Virtual Reality environments. For a more thorough analysis, hardware mockups that are cost and time intensive are constructed and used. Additionally, mockups do not always correspond to the latest datasets, are inflexible, and cannot be easily modified. The goal of this work is to extend the current visual validations of functional aspects, and expose the user through the possibility to directly manipulate objects in the car's interior to a real-time, life-like experience. In the automotive industry, the two immersive display systems mainly used today are CAVE-like systems [CNSD93] and head-mounted display (HMD) virtual environments, each for applications with slightly different focuses.

Table 1.1: Product attribute categories that have to be validated with functional assessments, the question that has to be answered with these assessments and typical examples

Attribute	Question	Example
accessibility	Is it possible for the driver to comfortably access all car functions?	Accessibility of the air conditioning control panel
usability	Can the car interface be comfortably used in the purposed way?	Comfortable grasp of the door handle during opening and closing
object visibility	Are drivers belonging to typical size percentiles capable of seeing all relevant functions inside the car?	Visibility of the instrument cluster
object clearance	Can objects be placed at the desired storage locations?	Placement of a cup in a cup holder
surrounding visibility	Have drivers belonging to typical size percentiles the possibility to see relevant objects surrounding the car?	Visibility of traffic lights

1.2.1 Evaluation of Ergonomics Issues

Projection-based virtual reality systems like the CAVE allow developers to virtually experience car interiors considering ergonomic aspects. Only a mockup of the car's driver seat is in the CAVE (Figure 1.1). The users are seated and can use their body as a self-reference. In this work, a three-sided CAVE with a resolution of 1920 by 1920 per side is used. The three-sided CAVE spans nearly the whole human field of view when looking straight ahead. The visual quality and comfort of this system leads to high user acceptance. This three-sided CAVE system is initially used for ergonomic assessments. At this stage in assessment, a team of experts meets to discuss individual aspects of a car's design . Due to the static nature of today's applications, these aspects mainly concern *surrounding* or *object visibility*. During the ergonomic assessments, interactive

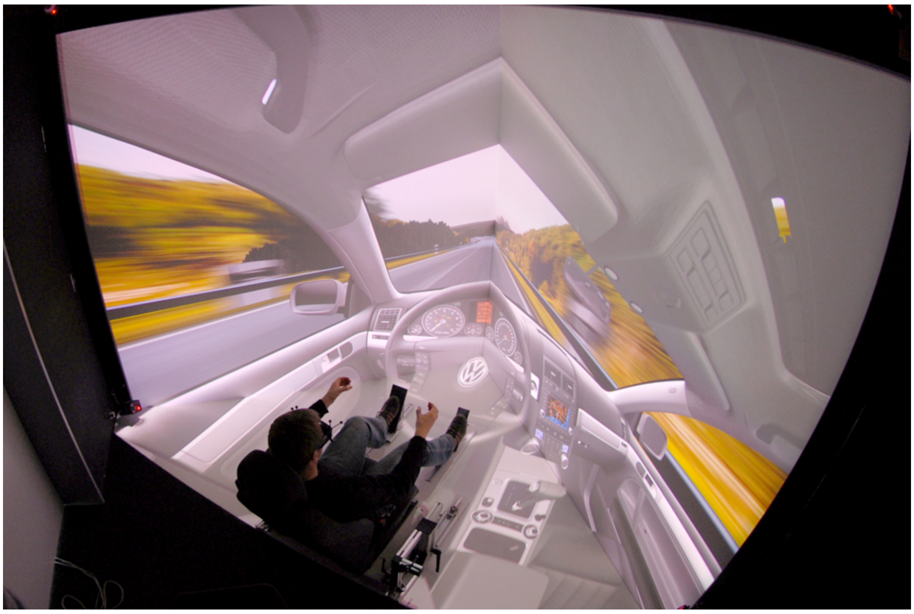

Figure 1.1: Typical CAVE application with a user evaluating a virtual car interior

parts of the car are not considered, or interactivity is reduced to a set of extreme positions or animations. Mainly the sun shields and the adjustable steering wheel are of interest for visual analysis, but any movable object in the car interior can cause occlusions. At this point, the *accessibility* and *usability* of car functions cannot be evaluated. Only a visual first impression is discussed by the experts.

Projection-based VR displays are challenging for direct finger-based interaction metaphors. Occlusion effects and focus shifts between users' real hands and the projection screens make it hard for some users to judge the location of their hands with respect to the virtual objects. Thus, the users' hands and fingers often penetrate the visual representations of the virtual objects, and this penetration is a major difficulty for robust finger-based interactions.

1.2.2 Evaluation of the Customers' Experience

Virtual Seating Bucks [SF08] (VSB) are an alternative to projection-based CAVE systems. In a VSB, a head-mounted display (HMD) is used to virtually complement minimal hardware mock-ups typically consisting of a steering wheel, a driver's seat, and floor pedals. HMD displays are capable of fully immersing the user in the experience by blocking the real environment and the user's own body from the perceived image (Figure 1.2). In this level of immersion judgment of hand-object relations is entirely based on the virtual hand representation. The drawback to HMD displays is their limited field of view, and the reduced comfort of the user when compared to a CAVE system. The HMD used in the scope of this work is a carefully set up Rockwell Collins SR80 providing 80° diagonal FOV with 100% overlap and SXGA resolution. This dis-

Figure 1.2: A two-user seating buck application for the assessment of electronic car interfaces [SF08]

play weighs 800 grams, and can be worn by the user for an extended period of time.

Virtual Seating Bucks are used to evaluate the potential customers' experience of novel car concepts. Various aspects of the car's characteristics are assessed by experts from the customer's point of view that compare the car's concept with past models and similar competitor models. The virtual assessment mainly considers *visibility* aspects. The car's functional properties are evaluated with the help of hardware mockups and tracked props. These mockups often represent deprecated datasets, and cannot be easily modified. The *usability* of the cars' multimedia interfaces and other functions, which play an important role in the customers' experience, cannot be evaluated within the car context. These aspects could be integrated into a Virtual Seating Buck application if it would provide a realistic interaction.

The evaluation of car ergonomics in the CAVE and the assessment of customer experience in a Virtual Seating Buck represent two important applications of the functional evaluation of car concepts in the product development process of the Volkswagen AG. They demonstrate the potential of direct finger-based interaction as a basis for the possibility to virtually assess human-car interaction in early development process phases.

1.2.3 Assembly Simulations

Virtual assembly simulations are the most accepted application area of immersive virtual technologies in industrial processes. The question to be answered with these applications is if a car part can be mounted at the desired location when certain assembly sequences are given, while also considering *object clearance* and *accessibility*. Currently, assembly sequences are validated by calculating possible assembly paths and by checking for collisions during an animation of the virtual objects along these paths. These virtual validations allow for the number of necessary hardware prototypes to be reduced. In addition, analysis of design variants is possible, and the ability to compare possible solutions with older car models.

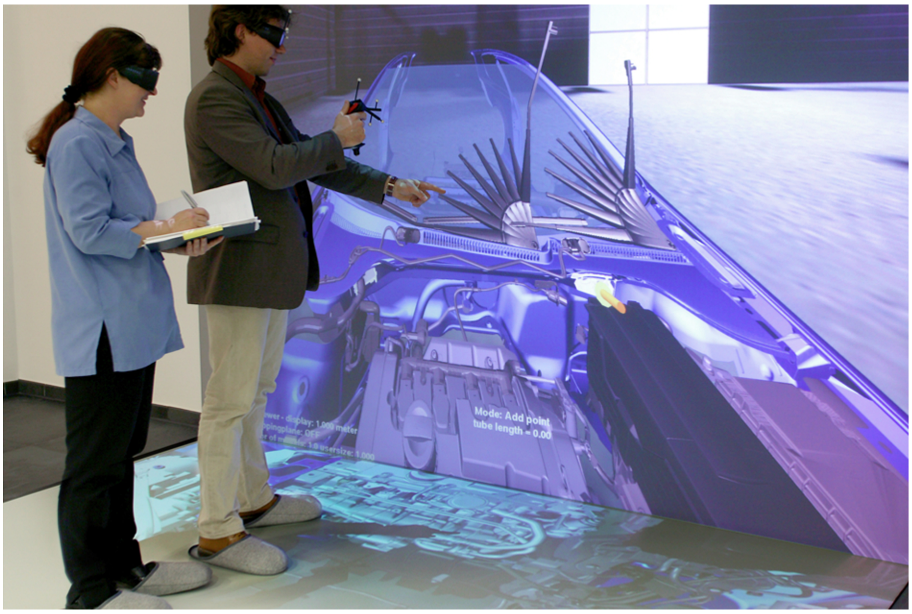

Figure 1.3: Interactive virtual assembly simulation with a user trying to mount a virtual car part in a L-projection environment

Special assembly challenges are evaluated within an immersive environment, and CAVE environments or L-Projections are used (Figure 1.3) in the evaluation. Challenging situations and possible solutions are presented and discussed in a more illustrative way with these mediums. CAVE and L-Projection environments provide the opportunity to conduct the assembly interactively. An interactive assembly simulation permits human factors to be considered, and allows experienced workers or assembly planners to incorporate their skills into the analysis.

Interactive assembly simulations usually rely on indirect interaction techniques based on controllers. Often a realistic assembly is not required but abstract metaphors can be used that provide a higher reliability. Direct finger-based interaction has the potential to strengthen the human factor of such assessments and to enable clearance analyses considering a mechanic's hands. Realistic assembly simulations enable the assessment of ergonomic issues during the actual assembly. The focus of the present work is on finger-based interaction in the car interior, but the interaction concepts presented are worthwhile and usable in other application areas as well.

1.3 Outline and Contributions

This thesis focuses on the development of robust and realistic direct finger-based interaction methods for the assessment of ergonomic aspects and functionalities of car interiors in CAVE-like systems and HMD-based virtual environments. The thesis is structured as follows.

Related Work First, a comprehensive overview of the related work is presented. This section deals with research concerning interaction metaphors and their classification, hand and finger tracking hardware, haptic feedback, and interaction evaluation.

Direct Finger-based Interaction The following chapter deals with the design, development, and implementation of direct finger-based interaction, and is based on two publications [MF10, MF11b]. This work is based on pseudophysical interaction techniques, which enable the realistic yet robust and efficient manipulation of virtual objects. Pseudophysical metaphors rely on reliable grasping heuristics with plausible object behavior. For direct finger-based interaction a well-defined set of simple grasping conditions enable the robust and reliable grasping of virtual objects in a realistic manner. A careful analysis of the function of typical features in the car interior revealed constraints that influence object reaction on user input. The geometrical description of these constraints and of the object reaction enables the implementation of plausible object behavior that is not physically correct but appears realistic.

Direct finger-based interaction in virtual environments has to deal with the complexities and irregularities of high-resolution 3D models. This thesis introduces the concept of Normal Proxies which significantly improves the robustness of virtual grasping actions for such scenarios. A further improvement of the grasping process is reached with the development of a refined finger-tracking system.

Tactile Feedback at the Finger Tips A novel tactile feedback system for the finger tips was designed, developed, and realized within this thesis. The system provides basic haptic feedback during interaction. This feedback is based on light-weight and unobtrusive pressure-generating actuators using shape memory alloys [SMF07]. As an alternative, vibro-tactile actuators were developed, which are mounted at the finger tips. The integration of these actuators into an optical finger tracking system is described here as well as the integration of the tactile feedback systems into the interaction framework.

Evaluation The proposed direct finger-based interaction techniques are evaluated and compared to a ray-based technique, which is the quasi-standard in industrial applications. An expert review shows that functional analyses of typical objects can be easily performed with the finger-based interaction technique. Furthermore, the review examines the overall usability of this approach. An extensive user study compares the proposed technique with conventional indirect interaction. For the success of direct finger-based interaction, realism is equally important as reliability and interaction efficiency. Because of this, the study combines quantitative measures for task performance with qualitative measures assessing subjective impressions such as interaction realism. And, the study evaluates the amount of user support provided by grasping feedback. Therefore, the proposed tactile feedback methods and visual feedback approach are compared to unsupported interaction. Parts of this chapter were published at IEEE Virtual Reality 2011 [MF11a].

Conclusions and Future Work The last chapter summarizes the contributions of this thesis, and reviews the research questions that were discussed in this introduction. Remaining aspects that are beyond the scope of this work are identified and a conclusion brings this thesis to a close.

Chapter 2

Related Work

A convincing interface to the virtual world is one of the key challenges for the implementation of innovative virtual methods and applications into industrial processes. Industrial applications especially have very high requirements concerning interaction robustness and interaction efficiency considering the fact that proposed innovations have to convince economically and because they have to compete with well established and consistently streamlined processes. However, an intuitive and reliable interaction is a prerequisite for the success of any application.

Consequently, the interaction with virtual environments has been a central research field within the virtual reality community since its inception. This chapter deals with past and recent approaches towards reliable and efficient interaction. Furthermore, it describes common techniques to transfer the users' hands into virtual environments, which is a prerequisite for direct finger-based interaction. Afterwards an overview of haptic devices and feedback shows how user support utilizing the haptic channel has been realized. Then previous work concerning the evaluation of interaction techniques is discussed. A description of the virtual reality framework, in which the proposed approaches are realized, completes the related work section.

2.1 Interaction Metaphors and their Classification

The way the users interact with virtual objects is one of the key characteristics of a Virtual Reality application and deserves special attention. The terms interaction metaphor and interaction technique have to be differentiated. Poupyrev et al. distinguished metaphors from techniques in a basic work on this topic [PWBI98]. They refer to interaction techniques as specific implementations of basic interaction metaphors. Consequently, there are just a few interaction metaphors that were further developed to specialized techniques that overcome shortcomings and challenges of the metaphors. Further developing these thoughts, Bowman et al. [BKLP04] classify interaction metaphors with respect to their task:

- *selection* — metaphors for the selection of desired objects
- *manipulation* — metaphors for changing object properties
- *navigation* — metaphors for the relocation of the users
- *system control* — metaphors for controlling functionality of the underlying (VR) system
- *symbolic input* — metaphors that are used to generate numeric or literal system input

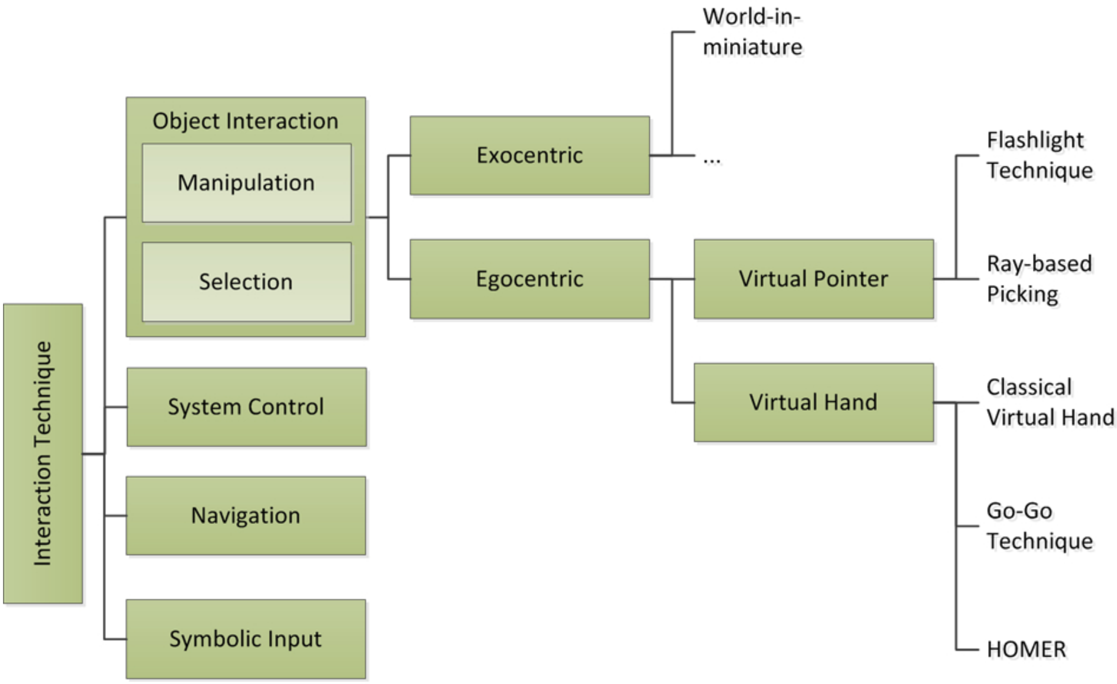

Figure 2.1: The basic taxonomy as it was suggested by Bowman et al. [BKLP04] and Poupyrev et al. [PWBI98]

For virtual audits of human-car interaction as they are targeted with this work, selection and manipulation prove most relevant considering users have to deal directly with parts of the car. Together the selection and manipulation metaphors describe the way we interact with virtual objects, where selection means the actual choosing of the desired object and manipulation means the changing of object properties. The object property to be changed in the applications focused on by this work is the location of the object. That is why in this case manipulation actually means the displacement of an object.

Often selection and manipulation are considered to belong to one interaction metaphor because in only few cases one makes sense without the other. Together they are further referenced as *object interaction* metaphors. Navigation, although of great interest to the VR research community, plays a minor role in the intended applications, where the users are seated on a mockup as the driver with a virtual car around them. System control and symbolic input metaphors are not the focus of this work either.

Poupyrev et al. and Bowman et al. distinguish object interaction metaphors to be *exocentric* or *egocentric*. The difference refers to the view of the user on the virtual environment. Exocentric metaphors provide an external view giving the user the chance to reference all objects in the virtual world, while with egocentric metaphors the users interact from inside the environment. They describe several exocentric techniques that are not the focus of this work. The authors define two basic egocentric metaphors, *Virtual Hand* and *Virtual Pointer*, both being further developed to specific interaction techniques. This basic classification does not consider the realism of the metaphors and it does not refer to the use of input devices. Consequently, very realistic interaction techniques with a direct mapping of the users' fingers to the virtual hand model are not differentiated from rough Virtual Hand techniques using a simple hand model as a cursor of an input device.

For a more detailed view on realistic hand-based interaction, as it is required for this work, and for the integration of more recent research approaches, adjustments to this taxonomy have to be made (Figure 2.2). First, a further distinction of egocentric metaphors into *indirect* or *direct interaction* needs to be added to the taxonomy. Indirect interaction is commonly used in industrial applications. An input device controls a cursor, which is then used to manipulate the scene. That is why indirect metaphors often are referred to as *controller-based metaphors*.

The division into indirect and direct interaction cannot be found in earlier work [BKLP04, PWBI98], but was introduced to differentiate interaction metaphors with respect to the relation between the users and the manipulated objects. Indirect interaction has the advantage in that an input device offers clear feedback of the interaction state and its changes. The disadvantage is that these metaphors do not allow an analysis of the corresponding real-world interaction itself. Wherever human-car interaction is the focus (e.g. in the analysis of ergonomics issues), direct manual manipulation is the only choice. However, direct manipulation of virtual objects suffers from a variety of limitations mainly in projection-based immersive environments and it suffers from a lack of haptic feedback. This missing feedback makes it necessary for the users to completely rely on their proprioceptive and visual sense to judge the relation of their hands to the virtual objects. However, in projection-based immersive environments, the user's real hands are located between the user and the display. Thus the real hands occlude virtual objects, but the virtual world is not able to occlude the real hands. In addition, the user's eyes focus in some cases on the real hand and not the display. Both problems affect the stereoscopic perception and limit the ease-of-use of direct interaction techniques. Despite these limitations, this very direct form of interaction is quite appealing and, in a variety of cases, serves as the desired option due to its direct correspondence to reality.

The introduction of indirect and direct interaction make it necessary to divide the virtual hand metaphor into *indirect virtual hand* and *direct virtual hand*. This refers to the fact that some of the virtual hand techniques implement an indirect metaphor, where an input device controls a hand-like cursor and others directly match human input to a virtual hand model. Motivated by more recent research approaches, two new interaction metaphors — *pseudophysical interaction* and *physical simulation-based interaction* are added to the taxonomy. The three direct interaction metaphors differ in the grade of realism they provide, while direct virtual hand is the most abstract metaphor, physical simulation-based interaction simulates the physical processes of the real world for object interaction.

The following chapters explain techniques implementing these basic interaction metaphors. The first chapter deals with the more abstract metaphors not incorporating a hand model. The second section describes hand-based metaphors including indirect and direct virtual hand and the new metaphors. For direct interaction metaphors, where the users manipulate objects directly with their hands, the selection part of the metaphors is considered as grasping. Grasping has been

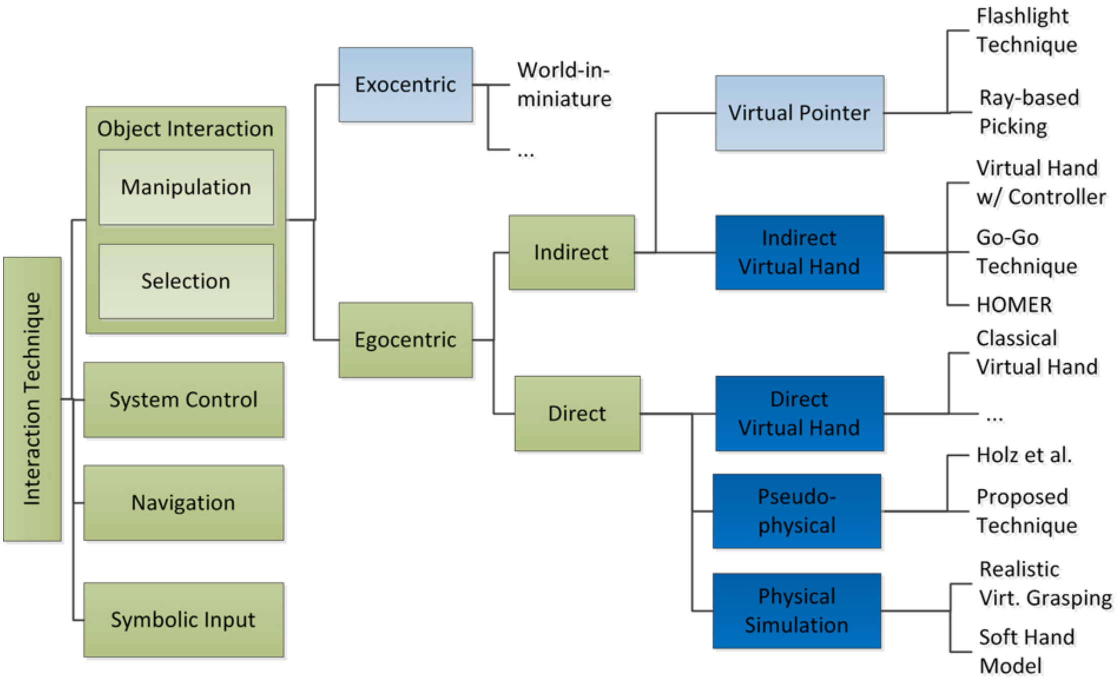

Figure 2.2: The extended taxonomy of interaction metaphors considering input devices and the realism of the metaphor. Some of the metaphors are more abstract (light blue) and some use virtual hand models (dark blue).

extensively studied in several research areas, such as robotics and character animation. These grasping approaches are not necessarily related to the direct manipulation of virtual objects but can provide techniques and methods that are inspiring for the realization of such techniques. That is why they are discussed in a separate section that presents some of these approaches and their applicability for direct interaction metaphors.

2.1.1 Abstract Interaction Metaphors

Exocentric and indirect metaphors are abstract metaphors that often are inspired by real human behavior or tools but do not have a direct paragon in real human object interaction. Many abstract object interaction techniques for virtual environments have been developed in the past for various virtual applications with specific interaction requirements. Bowman et al. [BKLP04] describe the most important approaches. Often these abstract techniques provide a very reliable and efficient way to select and manipulate virtual objects. As mentioned before, interaction metaphors can be distinguished to be exocentric or egocentric, according to the point of view of the user. The most common exocentric technique is World-in-miniature [SCP95] where the users hold a small copy of the virtual environment in their hands.

Egocentric metaphors give users the possibility to act from within the environment. An ego-

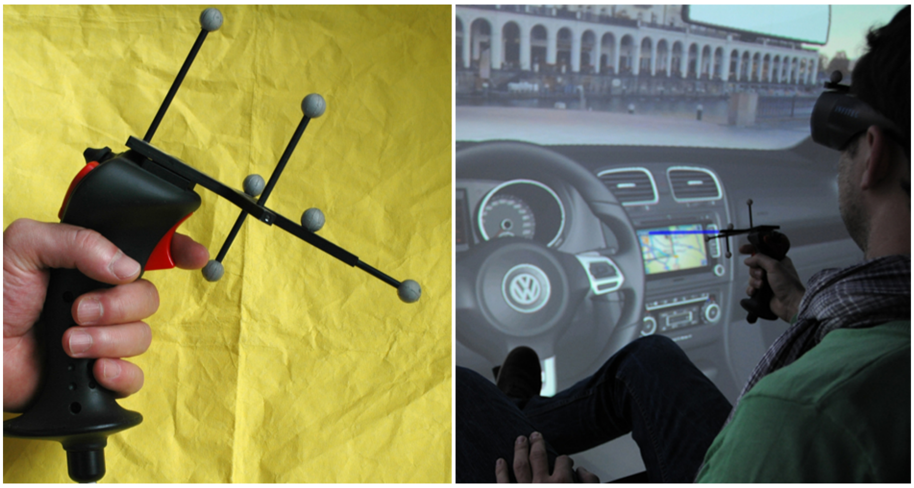

Figure 2.3: A common input device for immersive industrial VR applications — the Flystick (left) and the common implemented interaction technique — ray-based picking (right)

centric indirect metaphor is virtual pointer that uses the abstract concept of a virtual pointer geometry that is influenced by an input device. A virtual pointer technique that is widely used especially in industrial applications is ray-based picking (Figure 2.3). The tracked motion of a handheld 6D sensor is mapped on a three-dimensional cursor in the virtual environment, often visualized as an arrow or a simple hand model. The user points with a ray attached to this cursor at the desired virtual object and uses a trigger to confirm the selection. This trigger can be a button or a voice command. Usually a special device incorporating the 6D sensor and a number of buttons for triggering events is used, as for example the Flystick by the A.R.T. GmbH [Web10a] (Figure 2.3). Once selected, the object is attached to the ray and can be moved by the user. This technique is well evaluated and proves to be very efficient and easy-to-use especially when objects near the user are manipulated [BKLP04]. More sophisticated selection and manipulation tasks have inspired researchers to enhance this technique, as through widening the ray to a conic volume with the flashlight technique [LG94] or with a ray that automatically snaps to virtual objects for easier selection [SRH04, RHWF06].

As mentioned previously, the disadvantage of indirect and exocentric interaction metaphors is that the evaluation of the interaction itself and the transfer of interaction results into reality are not possible due to their abstract character. Consequently, for the applications focused on within this work, direct manipulation with the user's hand is essential.

2.1.2 Hand-based Interaction Metaphors

Common implementations of hand-based user interaction use the classical virtual hand technique [BKLP04]. This very simple form of a hand-based interaction lacks of realism and is not sufficient for human-car interaction analyzes. It does not implement thorough grasping strate-

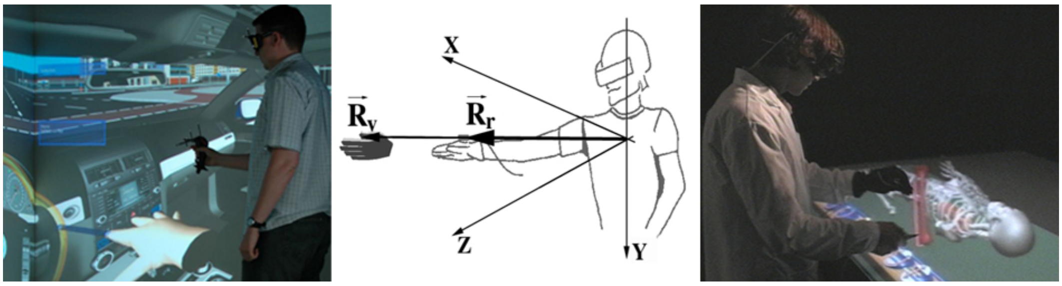

Figure 2.4: The indirect virtual hand metaphor combines controller-based interaction with a virtual hand model (left), the Go-Go-Technique tackles the range problem of virtual hand metaphors [PBWI96] (middle) and Cutler et al. combined a Fakespace PinchGlove™ with tracking for direct interaction at the Responsive Workbench [CFH97] (right).

gies based on fine-grain finger motion, but it combines hand collision with a button press for selection and derives manipulation from the pose of the whole hand. Thus the process of grasping an object properly as in reality cannot be simulated by this selection strategy. Furthermore the object manipulation is influenced by very rough movements of the whole hand. Fine-grained finger motion is not considered. This unrealistic manipulation is not sufficient for the realistic operation of a car interior that often incorporates sophisticated controllers. Due to the missing realism and due to the necessity for an additional button to enable and disable collision selection, this technique often is used like an indirect method using an input device with the virtual hand rendered with an offset to the device (Figure 2.4). Nevertheless, a commonly stated disadvantage of traditional virtual hand is that objects being located outside the arm's reach cannot be selected. In the context of this work this is not a disadvantage of the interaction metaphor. On the contrary the realistic and limited range of interaction is a prerequisite to enable the analysis of the the accessibility of objects.

The virtual hand metaphor was extended by the Go-Go-Technique by Poupyrev et al. to overcome its range limitation [PBWI96]. Here the distance of the virtual hand to the controller can be manipulated which extends the interaction scope of the user without navigation (Figure 2.4). Distance manipulation usually is realized utilizing an input device. A similar idea was pursued by Bowman et al. with the HOMER technique (hand-centered object manipulation extending ray-casting) [BH97]. Here the object is selected by ray-casting but subsequently the virtual hand of the user is attached to the object allowing the direct — but distant — manipulation of the object. These common enhancements of the Virtual Hand metaphor focus at a range extension coming along with a further reduction of realism. Being an advantage for some applications this makes them unusable for functional assessments of the car interior. However, when realism is not the focus, these techniques can be very efficient and they are commonly used in virtual environments.

For common virtual hand implementations it is necessary to provide a trigger for selecting an object. As for other indirect techniques, this trigger can be provided by buttons, speech commands and gestures. All these technologies have been studied and implemented by VR researchers. Speech and gesture recognition systems still did not prevail due to the learning effort and disambiguaties of these systems. However, to provide precise and easy-to-use triggers with button presses additional input devices have to be provided to the user. A possibility to avoid such a device is the Fakespace Pinch Glove™, which uses cloth buttons to detect pinches between the user's fingers. Such a system was used for example for menu interaction, text input or navigation [BWC$^+$02]. Van de Pol et al. attached tracking sensors to some of the fingers of a Pinch Glove™ and used it for direct picking. They assumed an explicit pinch and the presence of the pinched fingers inside the bounding box for the selection of the desired object [vdPRHP99]. Cutler et al. realized two-handed interaction within the arm's reach by using two tracked Pinch Gloves™ [CFH97] (Figure 2.4). A combination of cloth buttons and a bend-sensing data glove were used by [LZ99] to enhance a variety of glove-based interaction techniques. However none of these approaches realized direct finger-based interaction similar to what we are used to in reality. Nevertheless, the idea of using a pinch-sensitive finger tracking device is further developed with this work to increase the robustness of such a realistic interaction technique.

Few research approaches deal with really direct interaction techniques that are based on the users' fingers. This direct finger-based interaction is demanded for realistic interaction with virtual objects. These approaches can be further classified into approaches providing *physical simulation-based interaction* and those that realize a *pseudophysical interaction*. Pseudophysical interaction combines grasping heuristics with plausible object motion calculated as a reaction to the input of the grasping hand.

The physical simulation of direct interaction processes is the ultimate solution given that hand and finger manipulations of objects are governed by physical laws. Systems for rigid body dynamics, which are capable of simulating object-finger interactions, can be stable and fast. Baraff offers a comprehensive introduction to this topic [Bar97]. Rigid Body Simulations are widely used for game physics and are optimized for a high number of objects interacting with each other. To achieve this in a stable and robust manner, it is the case that constraints of the objects play a minor role in the simulation. The use of rigid body simulations in user interaction has been demonstrated [FTB$^+$00], wherein object reaction on indirect user input was calculated with the help of a physics engine. Albrecht et al. [AHS03] developed anatomically-based hand models with bones, virtual muscles and realistic skinning based on physical laws. Their very realistic hand animations could be the basis for physically-based interaction but the necessary calculations are too expensive for real-time interaction. Another work by Wan et al. [WLGP04] realizes realistic grasping of standard objects based on a Voxmap Point Shell collision detection, but this is only sufficient for simple objects of minor size.

Borst and Indugula showed a very realistic direct interaction of one user's hand with one freely-moveable object at a time using a commercially available physics simulation [BI06]. A physical representation of the virtual hand was attached to the tracked hand by springs. The interaction between the hand representation and the virtual object was computed by the physics engine (Figure 2.5). Typical object constraints of car-related scenarios are usually provided by these simulations. Multi-user support comes for free as well, since all human input has to be handled as interfering forces applied to objects. However, physical simulations usually have

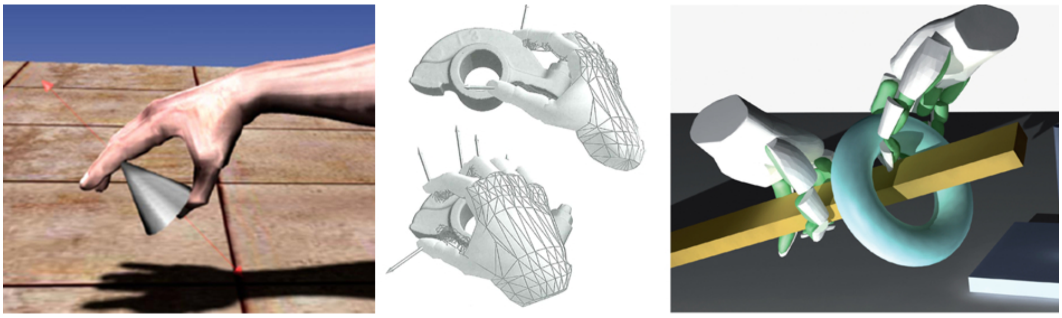

Figure 2.5: Holz et al. [HUWK08] proposed a pseudophysical interaction metaphor capable of handling a variety of freely movable objects (left), Borst and Induglia [BI06] realized very realistic virtual grasping using a physics simulation (middle) and Jacobs and Froehlich [JF11] enhanced their physical simulation-based grasping with a soft hand model increasing the friction applied to the virtual objects (right).

certain limitations with respect to the types of objects (e.g. only convex objects are supported). The present scenarios include arbitrary objects, such as thin sheets and very complex shapes. Physics simulations used for direct interaction remain challenging, especially when considering they have to compensate for instances in which the user interact with virtual objects in a way that simply cannot be solved using the simulation. Due to the absence of suitable output devices, it is not currently possible to appropriately restrict the users' motion with respect to constraints coming from the simulation. Therefore, the user can generate ultra-high forces applied to the objects and undefined object states that the physics simulation has to manage. Another promising approach by Duriez et al. [DCdlPAB08] accurately solves friction laws between the virtual hand and the object. They convey the practical use for non-immersive real-time interaction with simple objects, that however still relies on indirect controller-based selection and manipulation. Unfortunately, the integration of all aspects of direct finger-based interaction into a single application which is usable in the analysis of ergonomic issues in virtual cars has not yet been realized.

However, recent work shows promising advances towards a physical simulation-based interaction with virtual objects of various characteristics by combining rigid body dynamics with deformable pads attached to the virtual hand. These pads increase the interaction surface between fingers and objects (Figure 2.5) and thus they increase the friction and the forces that can be applied to the objects [JF11]. However, this approach still requires special data preparation and its robustness and ease-of-use remain unevaluated.

A pseudophysical interaction is realized when selection based on grasping heuristics is combined with a realistic object manipulation that is plausible enough to provide an object behavior similar to what can be encountered with real objects. Holz et al. described a technique that enables one-handed interactions with freely movable objects on a geometrical basis [HUWK08]. Here, all colliding fingers are defined as grasp pairs. In each frame the motion of the most significant finger pair is applied to the object, which in turn reproduces very realistic object behavior (Figure 2.5). For this interaction continuous finger-object collisions are essential. In the present applications mainly constrained objects are relevant. Continuous finger-object collisions cannot be guaranteed for these objects because they are not free to follow the fingers' motion. For interaction, only the most significant pair of phalangal contacts was taken into account, whereas multi-hand or multi-user interaction was neglected. Providing an interaction with multiple fingers (and hands and users) is necessary for the functional validation of virtual cars. Nevertheless pseudophysical metaphors are capable of enabling interaction techniques that are reliable and realistic enough to substitute and simulate the physical processes of user-object interplay during interaction. The present work describes an extended pseudophysical interaction technique that provides typical object interactions of multiple hands and users in a car interior.

2.1.3 Grasping

Direct interaction based on virtual grasping simulations has always been very popular despites its limitations. One reason for this is the powerful character of this metaphor. Moreover, in some cases problems can be avoided by the used display systems or a virtual hand representation with offset to the real hand could be used. Other important research fields that deal with human grasping are character animation and robotics. In character animation grasping simulations are used to animate realistic human grasping. In the robotics context, research is done to plan robotic grasps that are comparable to human grasps.

Due to Mas Sanso et al. [ST94] grasping simulations can be classified into *analytical* and *empirical* approaches. Analytical solutions calculate a physically correct position of the phalanges to form a valid grasp. Therefore a valid grasp is derived from the actual position of the hand. An appropriate grasping hand configuration is calculated and — under the avoidance of collisions — the virtual hand is relocated and reconfigured such that it matches the grasping configuration [ST94], [RBH$^+$95], [EKK$^+$07] and [BFH02]. Endo et al. use a hand model for their grasp planning that is configurable to match a broad range of human hands [EKK$^+$07]. Liu [Liu09] calculates a realistic and appealing grasping hand in motion from an initial grasping pose and the trajectory of the grasped object. Kry and Pai [KP03] derive realistic hand configurations from motion capturing data in combination with force sensors which obtain the grasping force while actually grasping real objects. The calculations which are necessary for analytical approaches usually prove too expensive to be conducted in real-time and therefore are mainly used for character animation and robotics where animations and grasp paths can be pre-calculated offline and simulation time is not as strenuously limited. Furthermore these applications do not have to take care of actual user hand input and can present a virtual hand having an arbitrary configuration disregarded to the users' real hands.

Empirical approaches observe human grasping scenarios and derive simple rules for grasping decisions. These grasping heuristics decide if the user has grasped a virtual object and they are generally easy to compute. Often grasping heuristics are used as a basis for analytical approaches. A taxonomy of human grasps helps to define rules for these heuristics. The observation of human grasping scenarios leads to just a few distinguishable grasps. A number of taxonomies that classify human grasps can be found in the literature, e.g. in [KI91], [US00] and [ZR01]. Some of them take both hands into account, some do not. They have in common

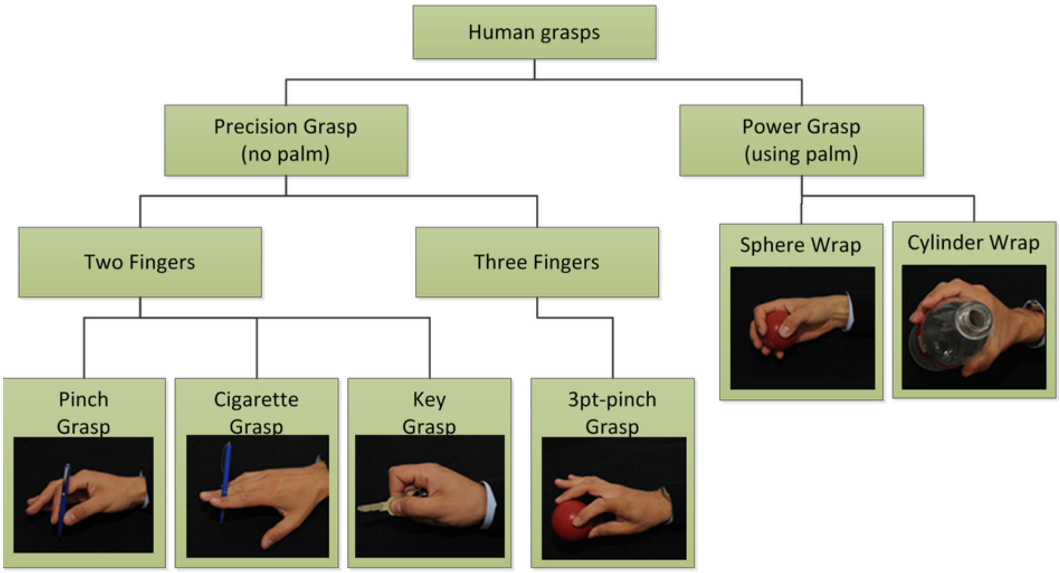

Figure 2.6: One-handed human grasps can be classified with respect to the involved hand segments.

that they mainly divide the grasps into two groups (Figure 2.6). Some separate them into power and precision grasps, some into grasps, using the palm, and grasps not using the palm. Most consider this segmentation to be equivalent, although Kang et al. discovered a difference [KI91]. The grasps differ in the special configuration of finger phalanges, hand palm and the grasped object. However they have in common that the virtual object is clamped by a number of phalanges of one hand or between segments of two hands. This clamping obviously is necessary to apply forces to the object preventing it from falling off the hand.

It seems that it is possible to derive simple rules leading to valid grasps when the situation of finger segments and virtual objects is analyzed. However, the conditions for grasp recognition have to be defined carefully. Rules that are too simple may allow for grasps that are not possible in reality. If the chosen conditions are too restrictive, it might be impossible to select objects. Empirical grasping heuristics do not calculate a plausible hand configuration that is necessary to avoid a penetration of the virtual object by the virtual hand.

Ullmann et al. [US00] define five rules for valid grasps with two hands. They introduce friction cones that simulate friction and analyze if an object is located between two fingers by calculating the angle between the contact normals of two colliding finger segments. The contact normal is the surface normal of the object at the collision location. These friction cones simulate the necessity that the finger segments can realize enough contact pressure to the object. Another interesting heuristic condition introduced here is derived from the observation of the distance of grasping-related phalanges. A grasp is released when this distance is increased significantly compared to the time of grasping start. With this rule the end of a grasp is detached from the collision situation and the opening of the hand for releasing is simulated heuristically. These powerful concepts have been modified and reused in other approaches and will be used in this work, too.

Another grasping-related aspect is the penetration of virtual objects by the virtual hand representation. As long as haptic feedback is not employed, the users' hands cannot be prevented from penetrating virtual objects during grasping. Burns et al. have shown that users more easily detect visual interpenetration than the visual-proprioceptive discrepancy that is introduced if a virtual hand representation is prevented from penetrating an object [BRP$^+$06]. In their study they utilized an HMD and did not consider interactive objects. It remains unclear if their findings apply to projection-based environments as well, where users are easily able to visually detect a dislocation of the virtual hand from the real hand. Moreover, it is not always obvious how a non-penetrating virtual hand should behave during complex interaction processes, when the users' hands are grasping through objects. For this work a co-located virtual hand is used to avoid a conflicting behavior of the real and virtual hand. Moreover the necessity of the presence of a virtual hand representation for interaction in projection-based environments is evaluated.

2.2 Glove-based Input

For direct interaction, the hands and fingers of the users have to be transferred into the virtual scene. Therefore, glove-based input devices combined with a tracking system or vision-based approaches should be used. Glove-based devices, like the commercially available CyberGlove™, use bend sensitive sensors to determine the configuration of the inter-phalangal joints of the hand. With a calibration procedure the length of the finger segments can be calculated. The pose of the outer-most phalanges can then be derived from the 6D pose of the hand target and a forward kinematic rotation and translation of the consecutive segments. That is why any sensoric error adds up and the overall error is biggest at the finger tips. Unfortunately the highest precision is needed there for finger-based interaction. Further disadvantages of this system such as poor results for the thumbs joint [KHW95] and its lack of comfort accrue.

Instead, for the present work, optical systems are preferred since they are able to precisely track the users' hand and finger motions. In particular, the finger tips are important considering that this is where the interaction takes place. The finger tracking system used here utilizes optical tracking to determine the position and orientation of the palm and the position of the tips of thumb, index and middle finger. These are the dominant and relevant fingers for interaction, ring finger and pinky can be omitted due to their minor contribution. For finger tip tracking sequentially flashing active markers are attached to the finger tips and a 6D target is mounted to the palm (Figure 4.1). A possible hand configuration matching the obeserved marker arrangement is calculated with inverse kinematics. Errors for the position of the finger tips may emerge only from the optical tracking system which is rather precise. The disadvantage is that the finger tracking is sensitive to line-of-sight problems that can occur when mockups are used or for seldom finger configurations such as a tight fist or joggle of the fingers of both hands. This finger tracking system is commercially available from the German company A.R.T. GmbH [Web10a] and works as a basis for all enhanced devices that are described in this thesis.

Recently the emergence of more and more markerless finger tracking approaches can be wit-

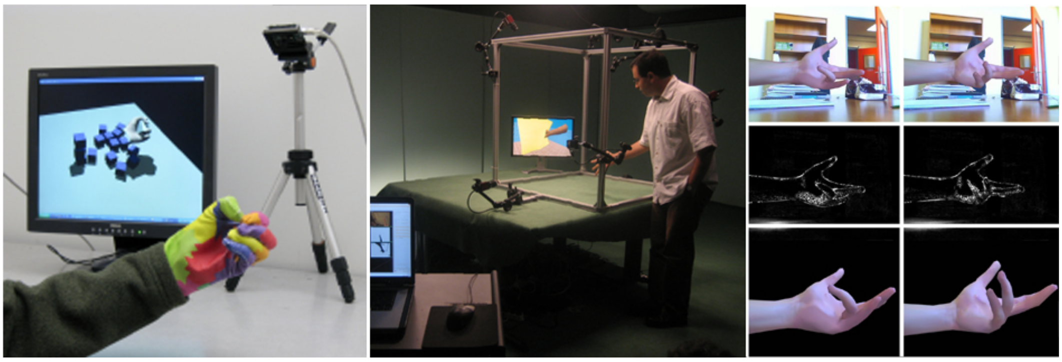

Figure 2.7: At the MIT a simple glove with printed colored patterns was developed that allows the estimation of a hand configuration from a single webcam image [WP09] (left), within the GrImage framework video-based hand tracking allows for a rough direct interaction with virtual objects [AMR^{+}07] (middle) and de La Gorce et al. estimated the hand configuration from a single camera image [dlGPF08] (right).

nessed. These approaches have the advantage that no hardware at the hands is disturbing the users during interaction. Especially in industrial applications it can be expected that the users' acceptance of any virtual application will increase with the reduction of obtrusive user mounted hardware. Wang and Popovic [WP09] introduced a simple glove with a printed colored pattern that makes it possible to derive a complete hand pose with a single webcam (Figure 2.7). Simple hand-based interaction has been shown to be possible with this approach. De La Gorce et al. also realized a hand pose estimation with a single camera image as input [dlGPF08]. Within the GrImage framework [Web11b] an application allowing bare hand interaction with physically behaving virtual objects has been realized [AMR^{+}07] (Figure 2.7). All these approaches have in common that they still do not provide the robustness and precision that is required for reliable direct interaction in complex and realistic virtual environments. Interaction space is limited and most approaches do not work with arbitrary backgrounds.

However, it can be expected that markerless hand and finger tracking will help virtual applications, relying on direct user input, to become more widely accepted in the nearer future. The latest developments on the gaming market make it probable that direct hand and finger input will become an important aspect of the mass market games of the nearer future. Cheap and reliable sensors using video and depth information like the Microsoft Kinect™ provide a significant advance for robust markerless finger tracking, as it was shown by MIT [Web11g]. Simple pointing gestures can be tracked very easily with such sensors and full hand gestures can be easily recognized. Once a markerless finger tracking approach is available that provides precision and reliability that is comparable with the optical tracking method used for this work, these approaches can be easily combined with the proposed interaction metaphor and will even increase the realism and naturalness of this metaphor.

Figure 2.8: The PHANToM force feedback system was the first commercially available system [Web11h] (left), the LWRIII of the DLR can be used for two-handed manipulation of virtual objects [HSA$^+$08] (right).

2.3 Haptic Feedback

During interaction in reality, people strongly benefit from the haptic feedback provided by real objects. It is difficult to provide this feedback in VEs and until now it is not possible to haptically render a virtual scene completely. However, several approaches that provide haptic feedback have been found and they are discussed in this chapter. Since solutions for haptic rendering still are not sufficient, it might be possible to use the haptic channel for additional information displayed to the user. Several approaches towards such a tactile feedback are also presented here.

2.3.1 Force Feedback

Haptic feedback can be differentiated into true *force feedback* and *tactile feedback*. Force feedback tries to simulate the real kinesthetic sensation for users during interaction with virtual objects. Both research approaches and commercial products are still of very limited applicability for direct interaction in projection-based virtual environments. The limited working volume and obtrusive design in combination with high installation and maintenance requirements [RHSM05] are the main critical issues.

However several researchers have developed devices that are capable of providing force feedback with a varying number of degrees of freedom. The most common way to render a virtual environment haptically is to use a force feedback robot. Research on this topic encompasses the haptic rendering as well as the development of devices that are capable of displaying forces generated by virtual objects to the user. Early devices like the PHANToM [MS94] were able to display the kinesthetic behavior of a 3D point in a virtual scene (Figure 2.8). More recent approaches like the commercially available Haption Virtuose™ $6D35 - 45$ [Web11d] or the Light-Weight Robot (LWRIII) of the German Space Research Agency [HSA$^+$08] are capable of

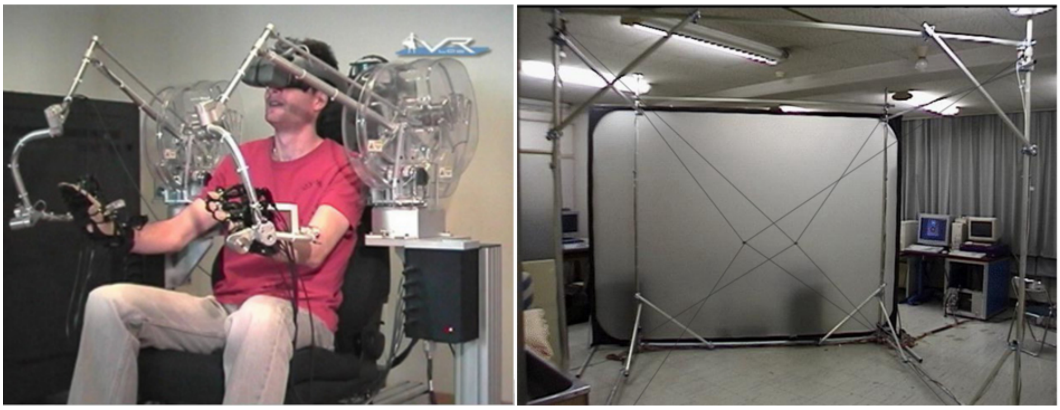

Figure 2.9: The commercially available Haptic Workstation by Immersion Corp. [Web11c] (left) and the Scaleable SPIDAR string-based force feedback system [BIS01] (right)

displaying 6D output including force and torque sensation (Figure 2.8).

For the haptic rendering of the virtual scene and the calculation of output forces, systems like the GOD-object algorithm [SZS95], the Voxmap Point-Shell algorithm [MPT99] or the Inner Sphere Tree algorithm [WZ09] are used. The necessary calculations for a complex car interior may impact real-time performance, since force feedback systems require fast update rates, usually at one kHz. Disregarded the quality of the haptic rendering and output itself, these devices are limited to indirect interaction metaphors because they provide a handle that can be influenced by the virtual forces. The robot arms disturb the users' view on the virtual scene in projection-based environments and limit the interaction volume.

These ground-referenced systems were combined with body referenced systems that directly limit the users degrees of freedom, to extend the realism of force feedback devices. For example the commercially available Haptic Workstation by CyberGlove Systems [Web11c] combines two of their CyberGrasps™ — which is a bend-sensing glove equipped with an exoskeleton — with two force feedback robots (Figure 2.9). The obtrusive design together with the calibration effort and the mechanical reliability of such a system impede the acceptance in real industrial processes.

As an alternative string-based haptic systems were developed that do not margin the users view and allow for bigger working volumes. The SPIDAR system evoluted from a single-point of contact system, over two-handed CAVE-sized systems providing three-dimensional force rendering, to a string-based system providing 3D force rendering for four fingers of each hand inside a cubic working volume [BIS01, WYHS02] (Figure 2.9). Although providing the possibility to enable the users to haptically feel a virtual environment, these devices suffer from a high complexity and are not appropriate to reduce common users' reservations about immersive virtual applications.

Figure 2.10: The commercially available CyberTouch™ [Web10b] (left), the tactile feedback glove incorporating voice-coil motors introduced by Pabon et al. [PSL+07] (middle), Inaba and Fujita suggested tactile feedback at the finger tips with a motor driven belt [IF06] (right).

2.3.2 Tactile Feedback

Tactile feedback uses the somatic senses of the skin as a channel to provide information to the user. A comprehensive overview of tactile devices is provided by Benali-Khoudja et al. [BKHAK04]. Tactile feedback has been used in a wide range of applications. Examples include the use of vests with integrated actuators as tactile displays [JNL04], sensory substitution devices for helping visually and acoustically impaired people [Bur96], and actuators mounted to input devices [RHSM05] for providing motion cues and collision information to the users of VR applications.

Usually vibro-tactile actuators are used because they are inexpensive off-the-shelf components that are small and reliable [Shi93]. The commercially available CyberTouch™ (Figure 2.10) device combines a conventional data glove with six vibration motors, one mounted to the upper side of the middle phalanx of each finger and one mounted to the inner side of the palm. The device can be used to create for example contact feedback with virtual objects. However, creating contact feedback at the back side of the finger phalanges does not seem to be the ideal location, since the contact happens typically at the finger tips. The weighty motors in combination with the tight fit of the glove create a relatively high pressure of the actuators on the skin of the fingers, which can lead to a transfer of the vibration output to the bone structure anticipating isolated stimulus perception [BB04].

In fact a limited amount of work exists that realizes vibro-tactile feedback at the finger tips. Kammermeier et al. designed a holistic kinesthetic and tactile display for the representation of object properties [KKHS04]. Pabon et al. incorporated voice-coil motors for vibrational feedback into a data glove without performing any evaluation [PSL+07] (Figure 2.10). Various other devices provide tactile information at the fingers for sensory substitution [Bur96]. As far as it is known, none of them was used for supporting the user during an interaction task.

Inaba and Fujita [IF06] propose a tactile feedback device concept quite close to the approach proposed later in this work (Figure 2.10). A belt wound around each fingertip is tightened by a motor to generate what they call a ”pseudo-force-feedback”. In contrast to the approach of the present work the whole fingertip is used as a display and no further discrimination of haptic patterns is possible. The actuators composed of a belt and a motor for each fingertip do not seem to be lightweight and unobtrusive enough to avoid disturbing the user during interaction. Aoki et al. [AMK$^+$09] have developed a similar feedback device but used a wire tightened by a motor for the feedback. Both works focus on the technical aspects and do not evaluate the influence of the feedback on direct interaction.

In general the interactional support of tactile feedback remains contentious [RHSM05, RK04]. The intention of this work is to develop a well-integrated tactile feedback system at the finger tips and to evaluate if it is capable of improving the objective and subjective interaction performance.

2.4 Evaluation

This thesis aims at the development of a realistic interaction metaphor that enables the functional validation of virtual car interiors. The success of this metaphor will be proven by a thorough evaluation of the developed techniques. Therefore several measures are considered. For a preliminary general evaluation, an expert review is conducted. Expert reviews are a technique coming from software development to evaluate the usability of human-computer interaction. Expert reviews usually reveal 75% of all usability problems with the help of only five experts [PRS02]. The potential of direct finger-based interaction for the focused applications and the overall usability of such a system will be shown with this method.

For a more significant evaluation of the techniques, more quantitative measures are needed. A common measure is task performance where certain tasks are conducted by subjects using different interaction metaphors and task completion times are compared as it is done, for example by Zhai [ZM98]. This measure is based on the assumption that a good interaction metaphor enables the user to fulfill a task in a shorter time. For direct finger-based interaction this measure is useful since an increase in grasping capabilities and grasping stability directly results in faster task completion. Therefore task performance is a measure of interaction reliability. The other focus of this thesis — interaction realism — cannot be measured by task completion times. Therefore subjective measures have to be used.

CAVE-like systems are used for the target applications as well as Virtual Seating Bucks employing a HMD. Although extensive work has been done in the fields of display technology and interaction metaphors, only a few studies exist regarding the impact of display choice on interaction. Comparing quantitative measures and task performance in several displays Boeman et al., Qi et al. and Demiralp et al. [BDF$^+$01, QIHM97, DJK$^+$06] report display-related differences in individual task performance and user satisfaction for search tasks and scientific data exploration. Kjeldskov evaluated virtual hand and virtual pointer techniques in a CAVE-like environment and a panoramic display [Kje01]. All studies identify display properties such as field of view, field of regard and image quality as a source for differences in task performance and subjective judgments. Consequently, differences in the usability of direct finger-based interaction have to be expected in the CAVE versus an HMD. The CAVE provides a higher field of view and it is generally more accepted by the target users. On the other hand, the HMD does not suffer from focus and convergence problems if direct hand-based interaction is required.

Furthermore the performance of direct finger-based interaction compared to traditional interaction metaphors is of interest. Although a very large number of interaction techniques have been developed, most of the work focuses on the use of large environments and thus on techniques enabling selection at-a-distance, such as ray-based picking or virtual hand enhancements like the Go-Go-Technique [PBWI96]. Other research examines and evaluates the particular characteristics of input devices [War90, ZM98]. Poupyrev et al., amongst others, compared virtual hand and virtual pointer metaphors within the arm's reach, which is quite similar to the scenario presented here [PWBI98]. However, their implementation of the direct metaphor does not provide realistic grasping. Instead, it resembles ray-based picking. Thus it is not surprising that both techniques perform comparably well within the reach of the user's hands. For this work, realistic grasping is generally desirable and the questions pertain to how well it performs compared to indirect interaction and how well it is accepted.

2.5 Software Environment

The direct finger-based interaction technique and the connection to all in- and output devices proposed within this thesis were implemented based on the VR framework Virtual Design 2. It was a commercially available VR system provided by the German company vrcom GmbH which is not on the market any more but was merged with the company icido GmbH [Web11e].

The rendering kernel was developed at the German research institute Fraunhofer IGD [Web11a] and is based on OpenGL. It is responsible for loading the geometries, building a hierarchical tree, and rendering the scene. The original VR system already was capable of handling input devices, such as the optical tracking including optical finger tracking, and it provided a powerful collision detection. Complex virtual environments, such as CAVE-like and HMD-based scenarios also can be realized with this system. Therefore PC-clusters are used, incorporating a master node with the main application and a number of clients running a rendering client and driving the displays these environments are built of. For the CAVE used in the context of this work, a cluster of six workstations was used. One of the works as the master and five are clients, each driving a pair of projectors. The HMD was driven by an additional client. The VR-system distributes all scene graph and camera changes of the master application to the clients.

Necessary extensions, such as the tactile feedback, the pinch-sensitive device and the interaction metaphor were implemented as dynamic shared objects (DSO) which are loaded at application start up time. The VR system allows several function types for the DSOs:

- *init functions* — executed once during the systems start
- *callback functions* — executed on previously defined events
- *loop functions* — called each frame
- *exit functions* — called at the systems termination

All initialization and cleanup issues necessary for the modules are handled within init and exit functions. The actual modules' functionality takes place in the loop and callback functions. The callback functions were used when the extension modules had to react on events of the VR system, such as collisions. Therefore a system-inherent interaction manager keeps track of all actions in the virtual environment such as the users' input. It is configured by a script which defines the static and dynamic properties of the scene. The description is based on the following primitives: events, actions and objects. Certain events evoke predefined actions, for example, a collision (event) of two geometries (objects) cause a callback function (action). Further functionality was implemented into loop functions that are called before each rendering frame.

2.6 Conclusion

This chapter summarized previous work related to this thesis. A lot of work has been done concerning reliable and robust interaction metaphors. However, most of these approaches lack of realism making it impossible to use them for functional audits of virtual car data. Direct interaction metaphors trying to provide realistic interaction with virtual objects have not been reliable and all-encompassing enough to be used in everyday industrial processes. Some of the aspects presented here can be used and refined towards reliable and realistic interaction.

In regards to the transfer of the users' hands into the virtual environment it is essential to state that until now it is not possible to refrain from hand-worn hardware for precise hand and finger tracking. However, this hardware has the potential to be enhanced to support the user during interaction. Extensive work has been done concerning haptic feedback. While true force feedback still is far away from being applicable in complex interaction scenarios, like those dealt with here, tactile feedback, using the haptic channel for information transfer, can be useful for enhancements of the grasping process. Until now mainly vibro-tactile feedback has been implemented and there is no realization of this feedback at the finger tips, so far. Few work exists that compares the performance of interaction metaphors when they are used in different display systems. The amount of user support induced by tactile feedback during interaction has not been proven so far.

Chapter 3

Direct Finger-based Interaction with Virtual Objects

A realistic interaction technique is mandatory for functional audits of virtual car interiors. At the same time this interaction has to be robust and reliable to be accepted in everyday industrial processes. This chapter describes the design and implementation of direct finger-based interaction as a robust and reliable yet ´highly realistic interaction technique. Reliable and robust selection can be achieved with appropriate grasping heuristics. The literature research proposed a set of grasping rules that are refined to reflect typical grasps found in the car interior.

Grasping in virtual environments is challenging for some users. Moreover some of the objects in the car interior have challenging characteristics for grasping. However, the proposed interaction metaphor has to provide reliable and realistic interaction in all kinds of virtual environments typically used in automotive development processes and with all objects found in the car interior. Therefore the grasping heuristics are enhanced by specialized grasping proxies and specialized hand hardware supporting the user with grasping challenging objects.

The object reaction on user input is described as pseudophysical calculations providing realistic object behavior. For this behavior various kinds of constraints that can be found in the targeted scenarios have to be considered. A classification of these constraints and their implementation is described in this chapter. The plausible and reliable interaction technique of direct finger-based interaction was described in two publications [MF10, MF11b].

3.1 Grasping Heuristics

The ability to grasp objects is essential for the realization of direct finger-based interaction in virtual environments. The moment of establishing a valid grasp is characterized as the selection of the virtual object. While grasping, the users' hand and finger movements manipulate the grasped objects by relocating them. The end of interaction is defined by deselecting objects through releasing the grasp.

As mentioned before, grasping has been extensively studied within several research fields. One way to detect human grasps is by using grasping heuristics. Here, valid grasps are detected by rules or conditions that define the beginning and the end of a particular grasp. Two parties are involved in a grasp: the virtual objects and the users. Grasping conditions are derived from the relationship of one party to the other. Therefore it is necessary to have a virtual representation of at least the users' hands in the VE. The basis of the heuristics developed here is the detection of collisions between hand and object geometries and their orientation to one another.

The definition of start conditions is subject to a trade-off. On the one hand, every single grasp intention of users has to be detected. On the other hand, unintended grasps have to be avoided as much as possible. Both requirements cannot be offended without annoying the users. The same is true for stop conditions. They have to provide stable grasping despite unintended hand movements, tracking jitter or even tracking interruptions. At the same time, it is unacceptable if a dropped object would remain sticking to the users' hands.

3.1.1 Common Grasps in Automotive Applications

The definition of start and stop conditions for the grasping heuristics was preceded by a thorough analysis of typical grasps used during interaction in a car interior. Therefore three users were video-taped while they used characteristic objects in the car. The grasps used during this interaction were classified with respect to common grasp classifications as they were introduced in Chapter 2.1.3. Only a limited set of distinctive grasps is used in these scenarios. Most of the objects are clamped between the fingers with variants of a pinch grasp. Bigger objects like the steering wheel or the door handle are picked with a cylinder wrap. While both hands are separately used for objects located on the respective sides of the driver, only few objects are picked with both hands simultaneously. Usually this is only done with the steering wheel. However the rules of grasping have to provide this possibility to reproduce the native human interaction. Especially for assembly simulation applications, two handed interaction is essential. Examples for common grasps used in the car can be seen in Figure 3.1.

One special case, that occurs quite often in the car interior, is push interaction. Here objects are not really grasped but pushed with a finger or the whole hand. Examples include various buttons and switches that only are pushed with a finger and the door or the sun shields that are either grasped or pushed with the whole hand. The direct finger-based interaction technique has to provide both, proper grasping of objects and their dislocation by simply pushing them. Especially for the door both interaction types — grasping and pushing — have to be provided considering that it is usually pushed to open and grasped to close.

3.1.2 Start Conditions

The grasping heuristics developed in this work are triggered by collisions between the finger phalanges and virtual objects. These collisions are detected on a per-triangle basis by the VR-Software. Whenever at least two finger phalanges collide with a single virtual object, both phalanges are checked if they establish a *grasping pair* with respect to this object. Valid grasping pairs are finger phalanges that have a virtual object between them. They are valid if the rules of a pseudophysical replacement of friction are satisfied. This means that physically correct friction is approximated by a geometrical representation called friction cones, as suggested by Holz et al. [HUWK08].

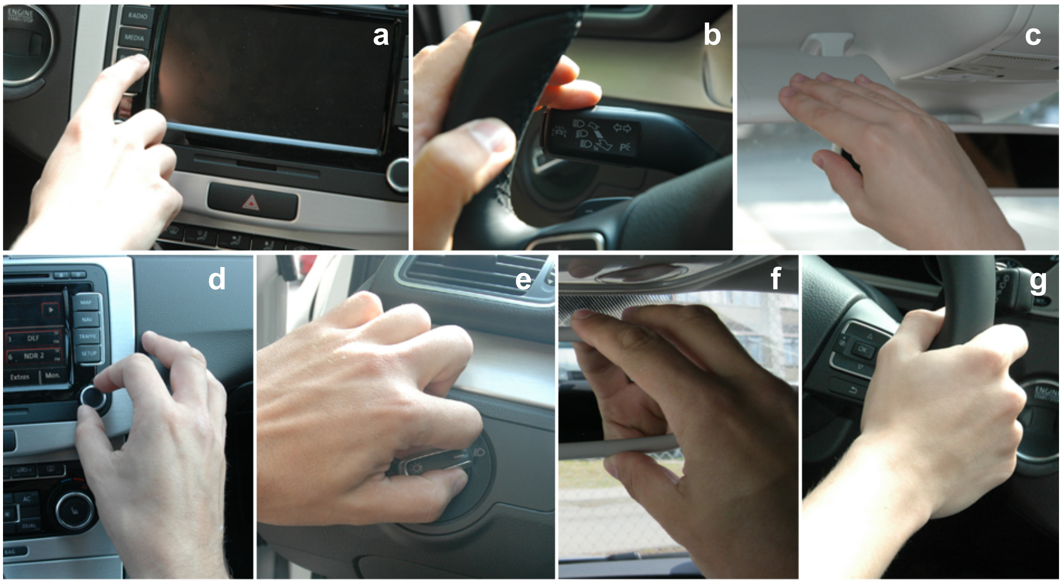

Figure 3.1: Common grasps in the car interior: pushing with a finger (a and b), whole hand pushing (c), tipp pinch (d), key grasp (e), three-point tipp pinch (f), cylinder wrap (g)

To explain the friction cone method, first several terms have to be defined. A *collision pair* consists of a finger phalanx and a virtual object that are colliding (in contrast to a grasping pair that consists of two collision pairs clamping the very same virtual object). The calculation of the *collision point* and the *collision normal* considers all triangles of the object's surface, which are involved in the collision. The collision point is defined as the center of all vertices of these triangles. The collision normal is calculated by averaging all involved face normals. Holz et al. suggest to collect all face normals within a certain radius of the collision center and to average them to calculate an appropriate collision normal [HUWK08]. Their approach smoothens rough surfaces and makes them graspable.

However, here another way of approximating normals has to be found since the car interior knows several box-shaped objects, including sun visors, the interior mirror or items that have to be placed in storage compartments such as books. Averaging normals of perpendicular faces leads to unusable collision normals that could prevent valid grasp pairs. The very same problem occurs while interacting with tiny objects that are not significantly bigger than the geometry of the finger tip. Here, a collision with the object can also include perpendicular or even opposing faces resulting in unusable normals if the averaging approach is applied. To avoid distortions of collision normals and the resulting problems with the grasping heuristics, appropriate *grasping proxies* are used that replace rough surfaces or perpendicular faces. These proxies are generated such that they provide normals that are compatible with the grasping heuristics. Details are explained further in section 3.1.7.

In order to calculate the friction cone and to decide if two collision pairs define a grasping pair,

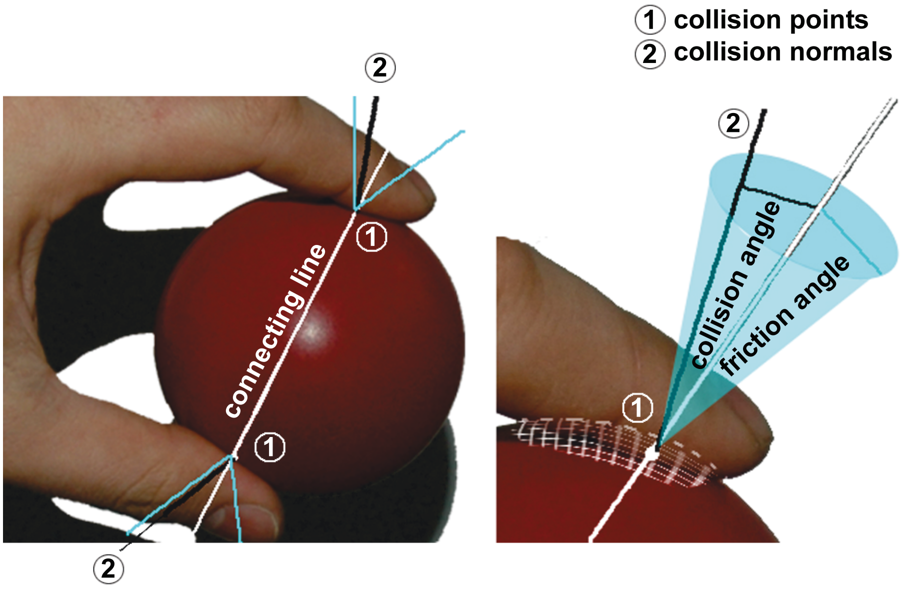

Figure 3.2: The elements of the heuristics start conditions: collision points and normals and the friction cone of a grasping pair.

the collision points $\overrightarrow{CP}$ are connected by a line $\overrightarrow{CL}$ (Equation 3.1). The angle α between the collision normal $\overrightarrow{CN}$ and the extended line is the *collision angle* (Figure 3.2) (Equation 3.2). The smaller the angle, the stronger the force that finger phalanges can apply to the object and thus the tighter the object can be clamped by the grasping pair. If the collision angles of both collision pairs are smaller than a pre-defined threshold — the *friction angle* — these collision pairs define a valid grasping pair. Surfaces with different friction values can be simulated by using friction cones with different angles. High thresholds simulate rough or even sticky surfaces, whereas small angles are used for smooth or slippery objects.

$$\begin{aligned}\overrightarrow{CL} &= \overrightarrow{CP(finger[1])} - \overrightarrow{CP(finger[0])}\\ \overrightarrow{CL'} &= \overrightarrow{CP(finger[0])} - \overrightarrow{CP(finger[1])}\end{aligned} \tag{3.1}$$

$$\begin{aligned}\cos\alpha(finger[0]) &= \frac{\overrightarrow{CL'} \circ \overrightarrow{CN(finger[0])}}{|\overrightarrow{CL'}| \cdot |\overrightarrow{CN(finger[0])}|}\\ \cos\alpha(finger[1]) &= \frac{\overrightarrow{CL} \circ \overrightarrow{CN(finger[1])}}{|\overrightarrow{CL}| \cdot |\overrightarrow{CN(finger[1])}|}\end{aligned} \tag{3.2}$$

An object is defined as being grasped if at least one valid grasping pair exists. A *grasp* is defined by a virtual object and all valid grasping pairs related to this object. It is irrelevant which particular hand the collision pair belongs to. Using this definition, multi-hand and even multi-user grasps can be uniformly treated. Moreover all relevant types of grasps can be ascribed to these simple start conditions.

3.1.3 Stop Conditions

A grasp of an object is valid if at least one valid grasping pair is detected. In each application frame, the virtual environment is checked for newly valid grasping pairs that need to be added to their associated grasp. Each existing grasping pair is checked if it is still valid with respect to the stop conditions. A grasp is valid as long as at least one valid grasping pair is associated with it. Two stop conditions are defined that are responsible for detecting disappeared grasping pairs.

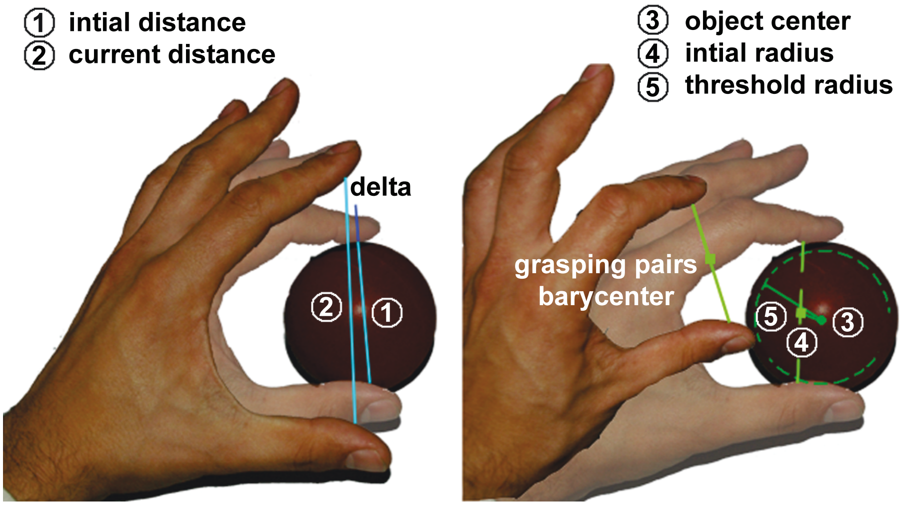

Figure 3.3: The stop conditions of the heuristics: contact point distance (left) and the grasping pair's distance to the object center (right).

The first stop condition compares the initial distance d_{t0} between the collision points $\overrightarrow{ColP_{t0}}$ of a newly detected grasping pair to the distance of the same points during interaction $\overrightarrow{CP_t}$ (Equation 3.4). A grasping pair is no longer valid if the current distance increased more than a pre-defined threshold from the initial distance. This stop condition aims at mimicking the users' intention to drop an object by opening their hand (Figure 3.3). The threshold is necessary to compensate for jitter and the imprecision of users trying to maintain a particular grasp. The calculation of the distance of the phalanges' collision points during interaction requires that the initial collision points are stored with respect to the local phalanx coordinate system of the involved phalanx (than referred to as $\overrightarrow{ColP'}$). In each frame the current pose of the involved phalanges M is used to transform the stored collision points to their current position, which is needed for the evaluation of the stop condition. These updated collision points are called the *contact point* $\overrightarrow{CP}$ of a collision pair (Equation 3.3).

$$\begin{aligned} \overrightarrow{ColP_{t0}}' &= \overrightarrow{ColP_{t0}} \cdot M_{t0}^{-1} \\ \overrightarrow{CP_t} &= \overrightarrow{ColP_{t0}}' \cdot M_t \end{aligned} \tag{3.3}$$

$$\begin{aligned} \overrightarrow{D_{t0}} &= \overrightarrow{ColP_{t0}(finger[1])} - \overrightarrow{ColP_{t0}(finger[0])} \\ \overrightarrow{D_t} &= \overrightarrow{CP_t(finger[1])} - \overrightarrow{CP_t(finger[0])} \\ d_{t0} &= \sqrt{\overrightarrow{D_{t0}} \bullet \overrightarrow{D_{t0}}} \\ d_t &= \sqrt{\overrightarrow{D_t} \bullet \overrightarrow{D_t}} \end{aligned} \tag{3.4}$$

$$\begin{aligned} \overrightarrow{BC} &= \overrightarrow{CP(finger[0])} + (\overrightarrow{CP(finger[1])} - \overrightarrow{CP(finger[0])})/2 \\ \overrightarrow{DD_{t0}} &= \overrightarrow{BC_{t0}} - \overrightarrow{OC_{t0}} \\ \overrightarrow{DD_{t1}} &= \overrightarrow{BC_{t1}} - \overrightarrow{OC_{t1}} \\ dd_{t0} &= \sqrt{\overrightarrow{DD_{t0}} \bullet \overrightarrow{DD_{t0}}} \\ dd_t &= \sqrt{\overrightarrow{DD_t} \bullet \overrightarrow{DD_t}} \end{aligned} \tag{3.5}$$

The second stop condition considers the users' intention to drop an object by moving their involved finger phalanges away from the grasped object. This is an important condition if constrained objects are involved, since they cannot always follow the hand movement. To better explain this condition, the *barycenter* $\overrightarrow{BC}$ of a grasping pair is defined to be the mean of both of its contact points $\overrightarrow{CP}$. The initial distance dd_{t0} of the grasping pair's barycenter to the object center $\overrightarrow{OC}$ is stored (Equation 3.5). This distance is continuously compared to the distance dd_t between the current barycenter and the object center. The grasping pair is valid as long as the current distance is smaller than the stored distance times a pre-defined threshold. In most cases it is reasonable to assume that the range of finger movements is small for small objects and grows commensurate with the size of an object, considering users tend to perform tiny and careful movements when dealing with small objects and coarser manipulations when dealing with larger objects. This expectation is considered by the threshold that is related to the object size.

If one of both stop conditions is fulfilled, the corresponding grasping pair is deleted. If no valid grasping pair remains for a grasp, the grasp is deleted as well. During interaction it is possible that grasping pairs appear and disappear. These stop conditions differ significantly from those described by Holz et al. [HUWK08] since they are not based on continuous collisions of the finger phalanges. Constrained objects cannot freely follow the movements of the finger phalanges. Thus, it cannot be assumed that the collision of the finger phalanges and the virtual objects can be continuously preserved during interaction. Instead, these stop conditions are heuristics, which try to match the users' intention to release an object.

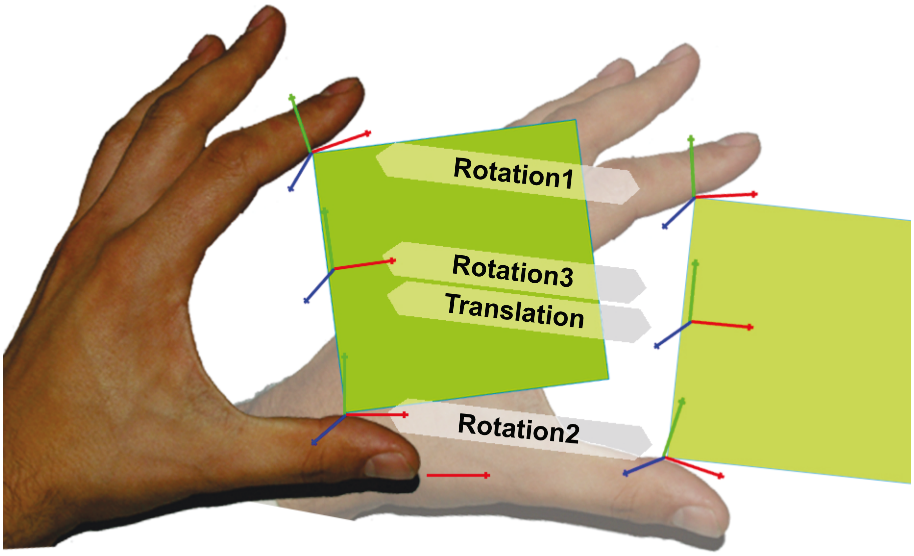

Figure 3.4: Transformations applied to the object: one translation and three averaged rotations.

3.1.4 User Input

During a valid grasp — from the moment the first grasping pair is detected to the moment the last grasping pair is deleted — the users are applying motions to the grasped objects. In reality, the forces applied to an object are responsible for its pose change. Since here no physics simulations are used, there has to be a plausible object motion derived from the motion of the finger phalanges.

For each grasping pair a frame-to-frame contribution is calculated based on the involved phalanges' movement. The contributions of all grasping pairs together comply with the user input intended by the motion of a grasping hand. The grasping pairs' translation — the so called *travel* ($\overrightarrow{S}$) — is simply derived from the movement of the barycenter $\overrightarrow{BC}$ of the grasping pair (Equations 3.6 and 3.7). For the rotation contribution, three different rotations are taken into account. The main rotation R_3 (see *Rotation* 3 in Figure 3.4) is derived from the movement of the line $\overrightarrow{CL}$ connecting the two contact points $\overrightarrow{CP}$ of the grasping pair. Therefore the angle α is calculated from the Dot Product of both vectors and the rotation axis $\overrightarrow{RA}$ is calculated from the Cross Product (Equation 3.10). The main rotation is caused by a translation of the fingers. Additionally, the rotation of each single phalanx has an effect on the object (see *Rotation* 1 and *Rotation* 2 in Figure 3.4). The rotations of the phalanges R_1 and R_2 can be directly calculated from the orientation change of each involved finger phalanx (Equations 3.8 and 3.9). These three rotations are averaged and constitute the rotational contribution of a grasping pair. For averaging the rotation contribution the SLERP algorithm is used [Sho85]. This algorithm was developed to animate rotational object motions but can also be used to compute the average of two or more rotation-describing quaternions (Equation 3.11). Therefore the rotation matrices R_1 and R_2 are transformed into quaternions (Q_1,Q_2) and a quaternion Q_3 is calculated from α and RA.

$$\overrightarrow{BC} = \overrightarrow{CP(finger[0])} + (\overrightarrow{CP(finger[0])} - \overrightarrow{CP(finger[1])}) * 0.5 \tag{3.6}$$

$$\overrightarrow{S} = \overrightarrow{BC_t} - \overrightarrow{BC_{t-1}} \tag{3.7}$$

$$R_1 = R(finger[0])_t \cdot R(finger[0])_{t-1}^{-1} \tag{3.8}$$

$$R_2 = R(finger[1])_t \cdot R(finger[1])_{t-1}^{-1} \tag{3.9}$$

$$\begin{aligned}
\overrightarrow{CL_{t-1}} &= \overrightarrow{CP(finger[0])_{t-1}} - \overrightarrow{CP(finger[1])_{t-1}} \\
\overrightarrow{CL_t} &= \overrightarrow{CP(finger[0])_t} - \overrightarrow{CP(finger[1])_t} \\
\cos\alpha &= \frac{\overrightarrow{CL_{t-1}} \bullet \overrightarrow{CL_t}}{|\overrightarrow{CL_{t-1}}| \cdot |\overrightarrow{CL_t}|} \\
\overrightarrow{RA} &= \overrightarrow{CL_{t-1}} \times \overrightarrow{CL_t}
\end{aligned} \tag{3.10}$$

$$Q_{all} = SLERP(SLERP(Q_1, Q_2, 1/2), Q_3, 1/3) \tag{3.11}$$

All translational and rotational contributions of all grasping pairs belonging to a grasp are applied to the grasped object. Therefore in each consecutive frame the following information is provided for each grasping pair:

- The translation contribution
- The rotation contribution calculated from three averaged partial rotations
- The location of the barycenter in the current frame
- The location of the barycenter in the frame before

3.1.5 Push Input

The grasping heuristics presented here enable the users to robustly grasp virtual objects. However, not every object in the car interior needs to be grasped for manipulation. Some objects, especially flaps and buttons, are usually pushed by one or more fingers or by the palm. This means that they are not clamped between the fingers but instead they back off from the finger phalanges in case of a collision while respecting their constraints. Since a pushable object always responds to any finger collision, this would impede robust grasping. Consequently, objects are then defined to be either pushable or graspable. Nevertheless, a combination of push and grasp input must be possible for large objects (Figure 3.5). In this case, parts of an object can be

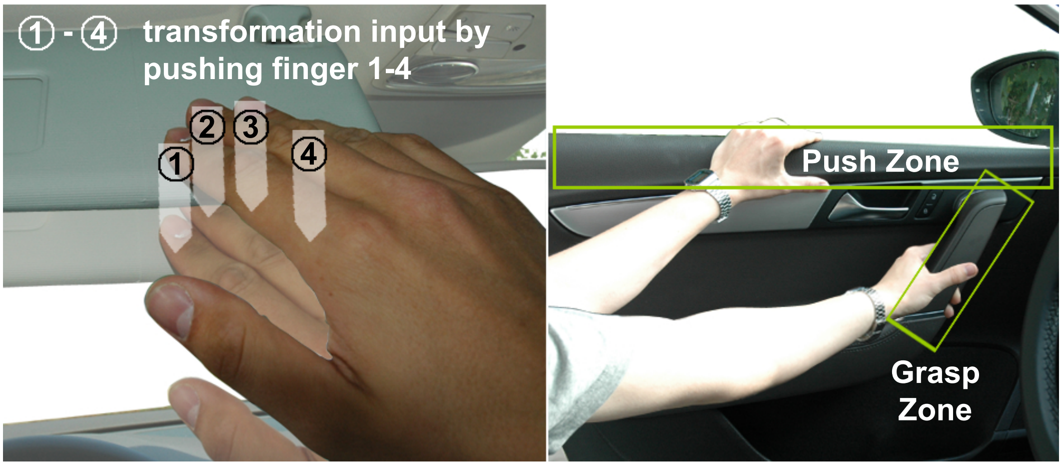

Figure 3.5: Translational input computed from the fingers' motion for push input (left) and push and grasp zones defined at the very same object (right).

defined to respond to push input and other parts can be defined to be graspable. Object reaction is then applied to the entire object.

To enable this intuitive interaction, a plausible object motion that results from pushing fingers has to be computed. The translational contribution of each involved collision pair is derived from the frame-to-frame distance that lead to a collision of the phalanx with the object. Rotational input is ignored (Figure 3.5). The contributions of all collision pairs are applied to the object as it is done with the grasping pairs for graspable objects.

This simple penalty method leads to object reactions that tend to be slightly exaggerated, since the exact finger penetration into the object is not calculated. However, this easily computable approach provides a convincing object motion because the back-off motion from an object is only minimally incorrect due to the limited hand velocity. Freely movable objects are not allowed to be pushed since their reaction on push input cannot be sufficiently simulated using this method.

3.1.6 Pinch-sensitive Grasping

It is sometimes hard for the users to fulfill the grasping conditions due to the size or shape of the virtual objects. It cannot be guaranteed that they are able to properly clamp the geometries of objects between their finger segments when objects are very thin or tiny. As an example the sun shields are only around $1cm$ thick. As another example the light switch has a handle that is approximately $1cm \times 1cm \times 3cm$. Especially with projection-based display systems, where the relation between fingers and objects have to be judged with an impaired visual channel, it is hard to correctly touch the upper and lower side of the objects with different fingers. Nevertheless proper grasping of such objects has to be realized to succeed with virtual functional validations.

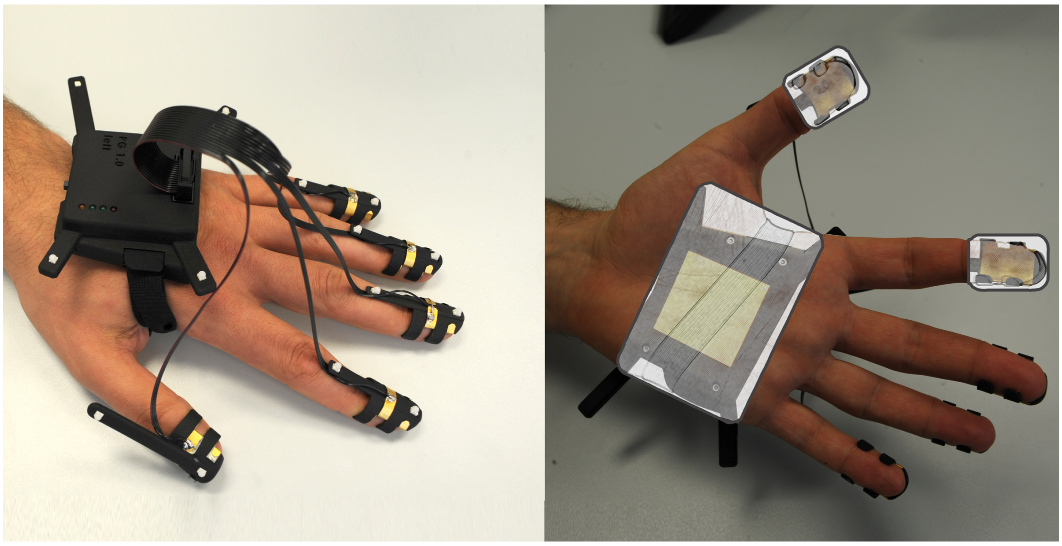

Figure 3.6: The pinch-sensitive finger tracking device (left) with an x-ray view of the incorporated conductive stripes (right).

It can be assumed that the users want to grasp an object when one or more fingers pinch with the thumb and simultaneously a collision between the virtual object and at least one of the involved finger segments can be detected. Especially for thin and tiny objects this finger pinch is a significant grasp condition. Moreover, users tend to pinch their fingers even when they grasp bigger objects due to the missing haptic feedback that would hinder the real fingers from sinking into virtual objects. Because of the significance of this pinch-based grasp condition, the selection of an object can be relieved from the primacy of correct finger object collisions and appropriate collision angles when this condition is fulfilled. Consequently, if the pinch information could be detected, this information could be used to improve grasping robustness.

Pinch-sensitive Finger Tracking Device

A possibility for the recognition of finger pinches was inspired by the Fakespace PinchGlove™. Therefore the optical finger tracking device was enhanced by a mechanism that is capable of recognizing pinching fingers. Considering average finger diameters are used for the virtual hand representation and due to the necessary finger thimble hardware it cannot be assumed that pinching fingers in reality always lead to pinching finger segments of the virtual hands. For a precise registration of the virtual hand with the finger tracking hardware a careful calibration step has to be conducted. Considering the everyday industrial use this calibration cannot always be guaranteed. With the proposed extension of the finger tracking device these real finger pinches can be detected reliably, disregarded the proximity of the virtual fingers. The pinch information can then be used to enhance and simplify grasp detection. The idea for such an extension was developed within the scope of this work. The conceptual and practical realization based on the conventional five finger finger tracking device was conducted by the A.R.T. GmbH together with

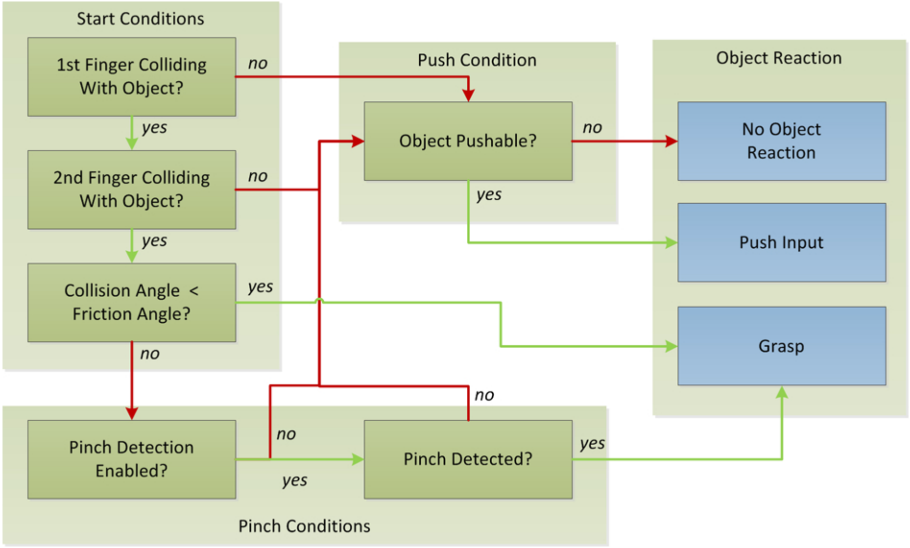

Figure 3.7: Diagramm of grasp decisions based on conventional start conditions, push input and pinch detection

the Bauhaus-Universität Weimar [Hae10].

For pinch detection, conductive stripes of Copper are incorporated into the finger thimbles and the $6D$ hand target housing (Figure 3.6). With them, low alternating-current with a voltage of $5V$ is introduced into the users' hands. When the fingers are pinched, circuits are closed with the human skin working as a resistor. Consequently, the voltage changes and this can be measured with a microcontroller. Thereby pinches of the real fingers can be detected immediately. Since every finger thimble is provided with a conductive stripe, pinches between the thumb and each of the other four fingers can be differentiated.

The voltage changes correlate with the strength of the finger press. Thus it is even possible to interpret the pinch states of each of the provided fingers as if they would be analogue buttons. The corresponding values are sent wirelessly using the Bluetooth protocol to an application communicating this information via UDP to arbitrary clients. This pinch information allows for an enhancement of the grasping heuristics. The pinch-sensitive system is used additionally to conventional grasp detection rules to improve grasp detection.

Enhancement of the Grasping Heuristics with Pinch Detection

According to the pinch grasps that were found to be common in the targeted applications (cf. Chapter 3.1.1), the grasping heuristics is provided with the pinch states of the following finger pairs of both hands: index finger-thumb and middle finger-thumb. These states are consistently updated with the pinch states reported by the control software of the pinch-sensitive device. This pinch information is used additionally to the start conditions presented earlier in this chapter. The complete grasp decision process is visualized in Figure 3.7. Pinch information is not used for

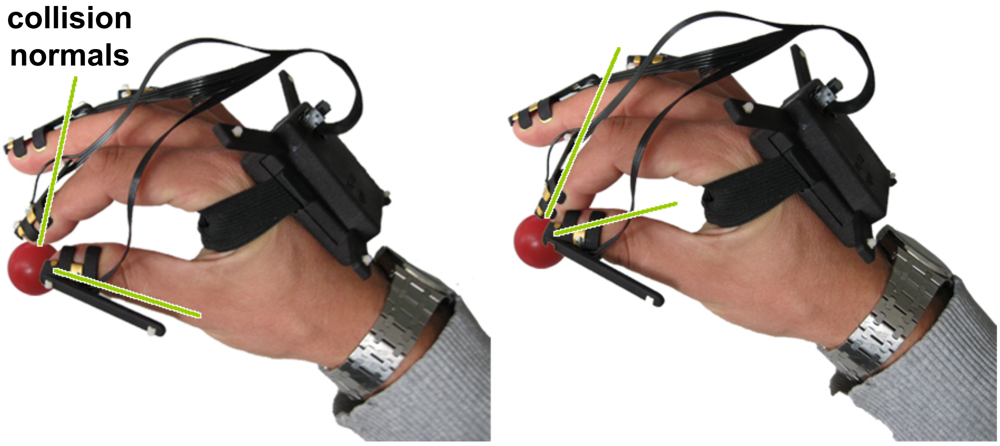

Figure 3.8: No grasp is detected due to the start conditions (left), with pinch detection the users' grasp intention can be satisfied disregarded non-opposing collision normals (right).

the decision on grasp release. For this decision the presented stop conditions are sufficient.

When the pinch-sensitive finger hardware is used, the start conditions proposed earlier are complemented with the following condition, in case of a negation of a valid grasp by the normal conditions. If a pinch-state is positive for one of the finger pairs and a collision of one of the two fingers with a virtual object is detected, a grasping pair is defined, containing the object and the two fingers. This grasping pair is handled as it is done with grasping pairs established due to normal start conditions. Consequently, multiple grasping pairs are possible when the users are performing a three-point pinch or grasp the object with two pinch grasps of the left and right hand.

With this enhancement proper grasping is detached from correct finger-object collision and appropriate collision normals in cases where the users pinch their fingers (Figure 3.8). By this, grasping robustness is improved especially for thin and tiny objects where proper collisions of fingers and objects cannot be guaranteed.

3.1.7 Grasping Proxies

The surface normals of objects are used to calculate friction cones, which determine the validity of a grasping pair. It is essential for robust grasp detection that the collision normal — which is the center axis of the friction cone — is a reasonable representation of the surface touched by the users. With wrong or falsified collision normals the grasping heuristics would not be able to decide correctly if the grasped object is properly clamped between the grasping finger segments. Collision normals are commonly calculated as the average of the surface normals of all faces, involved in the collision with the finger phalanges or of all faces located within a certain radius of the collision point [HUWK08].

There are three problems with the direct use of these collision normals in regards to the ro-

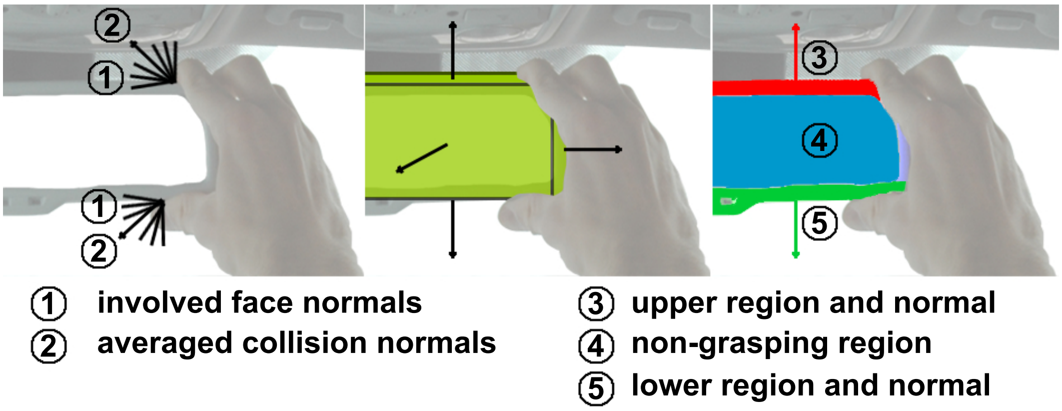

Figure 3.9: Averaging contact normals causes problems with the original geometry (left), collision proxies (middle) and normal proxies (right) provide reasonable normals

bustness of the grasping heuristic. First, there is the well-known problem of real world CAD models that still lacks a solution: flipped normals may occur for individual faces. Second, the collision of a finger with a box-shaped object (e.g. the interior mirror) can lead to a distorted collision normal due to the involvement of perpendicular faces (Figure 3.9). A similar problem occurs while interacting with an object that is as small as the finger geometry. It is almost impossible for the users to grasp such objects in a way that will result in a valid grasping pair, which requires that two fingers are colliding with faces having normals with almost opposite directions. Instead, for small objects it often happens that a single finger phalanx collides with a set of triangles that have perpendicular or even opposing normals. Third, automotive scenarios deal with complex geometry, which results in a high collision detection effort and may lead to decreased frame rates. However, this last problem has become less important due to efficient collision detection algorithms and faster processors.

A common solution for these problems is to substitute inappropriate geometry with simpler collision proxies. These substitutes are often used to approximate geometry in physically-based interaction scenarios [FTB$^+$00]. They can also be used with the approach proposed here to define appropriate normals for grasp detection. These simplified proxy geometries are only used for grasp detection and the computation of the physically realistic object motion, whereas the object is still rendered from its original representation.

The collision proxy approach performs collision detection through a simplification of the original object geometry. While this is an advantage concerning speed and grasp detection robustness, it introduces an error due to the simplification involved. This work therefore suggests to use the concept of *normal proxies* instead of collision proxies. Normal proxies extend the object description by additional normals, which are only used for grasp detection. The original normal responsible for lighting is not changed. With this approach, the collision detection still refers to the original geometry but for grasp detection appropriate proxy normals can be used. This allows for overcoming the annoying flipped normals problem, but more importantly, the problem of averaging unsuitable normals from a neighborhood of a colliding phalanx can be

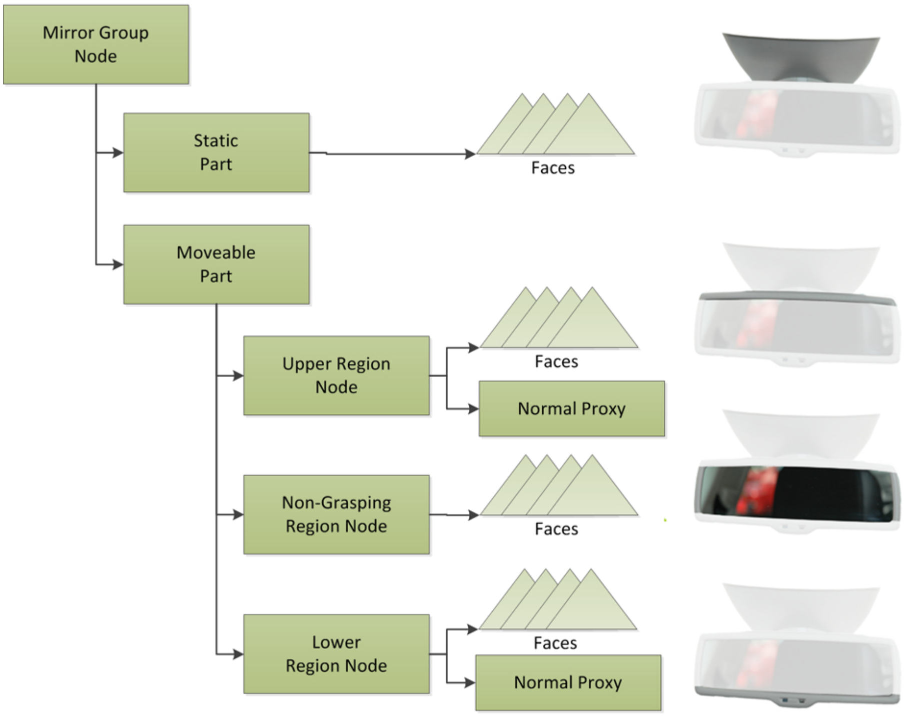

Figure 3.10: Segmentation into grasp regions with the scene graph hierarchy. The objects' node is divided into a static and a movable part. The movable part contains three regions: the up and down region, each with an appropriate normal proxy, and a non-grasping region. All transformations are applied to the movable group node.

overcome.

One challenge with normal proxies lies within data preparation. A normal, providing robust grasp detection, has to be assigned to each face of the object. The direction of the normal proxy depends on the location of the face on the object's surface and the way the user is grasping this particular object. In the scenarios used here, usually a small number of varying normals are necessary per object. All objects are designed to be used in a defined way and can be reduced to a small number of grasping surfaces or areas with a common normal proxy. To keep the data preparation effort low, a semi-automatic approach is used to define the normal proxies. Objects are divided into grasp regions which correspond to the areas where the fingers are typically located during a grasp. Each grasp region contains a set of faces and has a common normal proxy. Additionally, to avoid unintended grasps, regions are introduced, where an object cannot be grasped. The differentiation into grasp regions is realized with a specialized scene graph hierarchy where the node representing the whole interactive object holds a group node for each grasp region, containing all related geometry and a normal proxy object (Figure 3.10).

The concept of these different grasp regions is explained using the interior mirror as an example shown in Figure 3.9. The interior mirror has three grasp regions. The first region has a grasp normal that is pointing upwards and contains all triangles of the upper part of the mirror. The second region groups all faces of the lower part with a normal pointing downwards. The third grasp region contains all remaining faces and is defined as not graspable. This subdivision is due to the fact that the interior mirror is usually clamped between the index finger and thumb on its upper and lower sides. Without a proxy method, the averaging of normals of colliding phalanges may make it impossible to actually grasp the mirror. The normal proxy approach allows for a collision detection on the original geometry, but the grasping heuristics is provided with normals that provide a robust grasp (Figure 3.9). This means that disregarded to the actual location of the finger phalanx an appropriate normal is used for grasp decision when the touched face lies within a grasp region.

All three approaches (using the complete original geometry, using collision proxies or using normal proxies) are compatible with the grasping heuristics proposed here. Collision proxies are fast and robust but inaccurate with respect to the actual collision. Using the original geometry may make it impossible to grasp an object. Normal proxies are a good compromise for increasing the robustness of grasping heuristics. The disadvantage is that they cannot be easily constructed by a fully automatic approach. However, since car interiors are typically constructed from a set of defined building blocks, the normal proxies need to be created only once and can then be stored with the objects in the building block library.

3.1.8 Conclusion

The previous section described the design and implementation of refined grasping heuristics making it possible to reliably and realistically select and deselect virtual objects by grasping. Simple rules of grasping based on the collision of finger segments with the virtual object combined with a simplified friction simulation by friction cones enable robust and realistic grasp detection. Some objects do not allow to use the conventional friction cone method due to their characteristic surface and the resulting consequences for the surface normals. For these objects a simplification called normal proxies was introduced that ensure realistic and reliable grasping for all kind of objects. Especially the grasping of thin and tiny objects further benefits from the integration of pinch-sensitive finger tracking hardware that additionally considers the pinches of the users' fingers as grasp conditions. With these enhancements the proposed grasping heuristics provide the reliable and realistic grasping and releasing of virtual objects. Furthermore, this section described which information is retrieved from the users' hands and fingers motion under grasp condition. The next section describes how a realistic object reaction on this input can be calculated on a pseudophysical basis.

3.2 Object Reaction

In reality, human-object interaction is a result of physical processes. Basically these processes can be simulated for user interactions with virtual objects, but as it was explained before, until now the complex scenarios found in automotive applications cannot be simulated completely. One aspect of the physical processes is, that the users apply forces to the objects with their hands. The previous Section 3.1 described how these processes are translated into a pseudophysical model and how user input is calculated from the motion of grasp-related finger segments. The object's behavior is a reaction based on this input and should be physically plausible, as well.

Forces applied by the users interact with the inner forces of the object (e.g. gravity, friction and inertia) and cause it to move. A further influence on object motion can be caused by restrictions of its degrees of freedom (DOF). An object has a total of six DOF if it is not restricted (3 DOF of translation and 3 DOF of rotation). If an object has six DOF, the force that is applied to the object results in a motion in the direction of the superpositioned input forces (e.g. user input, gravity, etc.). For restricted objects having less than six degrees of freedom, the objects' reactions are more difficult to compute.

This chapter describes the calculations leading to a pseudophysical realistic object reaction on user input. First, common object types found in the car interior and their characteristics are analyzed. Then it is explained how objects that are not restricted react towards user input while being grasped. Finally, restricted objects are discussed, wherein constraints can occur and the objects' motion is influenced by these constraints.

3.2.1 Typical Object Characteristics

Most of the objects that can be found in the car interior have limited degrees of freedom. The objects can be classified due to these restrictions.

First there are *freely movable objects* that can be placed anywhere in the car. Strictly speaking, these objects are not part of the car interior itself but have to be considered when ergonomics issues are analyzed. Freely movable objects that usually can be found in a car in real life, include small objects, such as CD-covers and music players, and bigger objects, such as bottles or road atlases.

1R objects, that can be rotated around one axis, probably are the most common type of object that can be found in a car interior. However, they have to be further divided due to the way they are handled. One type of 1R objects are *flaps*. Flaps usually are rotated around a rotation axis, that is located on one side of the object, by a translation of the grasping hand at the other side of the object. This kind of rotation is further referred to as *lever-like rotation*. Examples for flaps are the sun shields or any kind of door including the driver's door or the glove compartment latch. Flaps usually are further restricted by stop positions they cannot be rotated beyond. Often this kind of object is not properly grasped but pushed by the hand or some fingers.

Another type of 1R objects are *controllers*. Controllers are small objects that are rotated around an axis that is going through the center of the object. Consequently, they are rotated by a rotation of the grasping hand enclosing the whole object. Since controllers usually provide a dedicated car functionality they are further restricted by stop positions or positions they snap into. One special case is the steering wheel. Due to its size and the way it is constructed, it can be rotated by a lever-like rotation of one hand as well as by a controller-like rotation of both hands.

There are only few objects that can be rotated around two axes. Examples for *2R objects* are levers that provide car functionality, such as the operation of the indicator lights or the wipers. They usually are pushed with one or more fingers in only one of the two rotational dimensions. This rotation is similar to what was defined as a lever-like rotation of 1R objects.

The only object that can be rotated freely is the interior mirror. Their are diverse ways to handle this *3R object*. The most common way of mirror adjustment is to grasp it with one hand at the most right hand side. The motion of the interior mirror is further restricted by the surrounding objects, the wind shield and the headliner.

There are only few objects that have constraints allowing for a translation of the object. One type of *1T objects* are *buttons*. They usually are pushed with one finger and perform a translation along one vector until they reach a stop position.

Some of the objects in the car interior follow simple kinematic chains involving a translational motion. One special case is a controller — the light switch — that can be rotated around an axis and moved along this axis to switch additional car functionality. Another example is the steering wheel adjustment. On the one hand the steering column can be rotated with a lever-like rotation around one axis, on the other hand it can be pushed or pulled along another vector. The resulting motion of the steering wheel mounted to the end of the column follows a superposition of both restrictions. Of course the steering wheel itself can be rotated around the steering axis, as well.

Generally, the motion reaction of objects on user input is strictly influenced by constraints restricting their degrees of freedom. There are several types of such constraints. These types and there influence on object motion are discussed in the following sections.

3.2.2 Unrestricted Objects

As mentioned before, unrestricted objects have six degrees of freedom. In reality, freely movable objects that are not grabbed by the users would fall due to gravity until they would hit the floor, for example. The simulation of gravity is not necessary for the scenarios dealt with here. Objects that can be grabbed and moved through space without any restrictions nevertheless are necessary (e.g. for clearance analyses). If a non-restricted object is grasped in reality it follows the user-induced forces instead of secondary forces including gravity and air resistance. This is the case because input forces are necessarily much higher than object inherent forces. The manipulating person consistently adjusts the input forces to counterbalance gravity and air resistance. That is why, while an object is properly grasped, only the user-induced input forces have to be considered to provide physically plausible object behavior.

The object's reaction to the user input introduced by a grasp can be easily calculated from the rotations and translations of the corresponding grasping pairs. The grasping heuristics provide the location of the grasping pairs' barycenters in the current frame and the frame before, the translation contribution and the averaged rotation contribution of the grasping pairs. The center of mass of the objects is irrelevant for this pseudophysical calculation of object reaction. Instead the center of rotation — the *grasp center* — for unrestricted object motion is computed as the mean of the last frames' barycenters of all involved grasping pairs. The rotation contributions of all grasping pairs are averaged with the SLERP algorithm and are directly applied as the object rotation. The object translation is defined by the averaged travels of the grasping pairs.

The object's transformation matrix representing the user intended object location and orientation is calculated as follows. First a transformation is calculated that moves the object from the last frames position — defined by the matrix M_{t-1} — in a way that the grasp center correlates with the world origin. Therefore a matrix representing the grasp center (M_{gc}) is inverted. Then object rotation is calculated as the matrix M_{rot} from the quaternion representing the averaged rotations of all grasping pairs. Object translation is represented by the matrix M_{trans} which only holds the averaged travels of all grasping pairs. Finally the object is moved to the desired location by translating it back by the matrix M_{gc}. The final object transformation $M_t$is calculated by the multiplication of all matrices in the presented order (Equation 3.12).

$$M_t \; = \; M_{t-1} \bullet M_{gc}^{-1} \bullet M_{rot} \bullet M_{trans} \bullet M_{gc} \tag{3.12}$$

This method of pseudophysical behavior addresses both, coarse motions defined by long distance movements of the whole hand and fine-grained motions as they are intended by tiny shifts of the two phalanges belonging to one grasping pair.

3.2.3 Objects with Constraints

Freely movable objects move as expected by the user. However, most objects found in automotive scenarios have restricted degrees of freedom since they are mounted to the car body in one way or another. Therefore three different types of constraints are defined:

1. Mounting constraints result from the way objects are mounted onto the car body. An example includes the glove compartment latch mounted to the cockpit by a hinge. These constraints are well known from the field of Rigid Body Dynamics.
2. Functional constraints further influence object motion within a mounting constraint and represent a certain function of the object (e.g. lock positions of controllers).
3. Location constraints define the clearance of an object. They are caused by collisions with other objects. Constraints of this type are often avoided by functional constraints. In this case, engineers restrict the objects' motion such, that they do not collide with other objects.

Each type of constraint has a certain influence on the objects' reaction to user input. The following sections explain how this influence is addressed by the direct finger-based interaction metaphor.

Mounting Constraints

Mounting constraints are well known from Rigid Body Dynamics, where they are generally considered as joints linking objects to each other. Joints are mainly characterized by the number of remaining degrees of freedom they provide. A list of typical mounting constraint types include the following, examples are given in Figure 3.11:

- *Hinge joint*: Rotation around one axis.
- *Revolute joint*: Rotation around one axis but this axis is located in the center of the object.
- *Saddle joint*: Two dimensional rotation, often realized as a kinematic row of two hinge joints.
- *Ball and socket joint*: Three dimensional rotation.
- *Prismatic joint*: One dimensional translation along one vector.
- *Cylindrical joint*: A revolute joint that additionally allows for a translation along the rotation axis.

Additionally there are objects that follow kinematic chains of two or more joints. Kinematics that can be found in a car interior are usually quite simple. For example the steering wheel adjustment is defined by a combination of a hinge and a guide plate resulting in a 1R1T constraint. These simple kinematic chains can be represented by the hierarchical structure of the scene graph. Each joint in the chain can be modeled by a transformation node arranged in a hierarchy similar to the kinematic chain. Object reaction with respect to user input is calculated and separately applied for each of the joints. The interaction object is child to all group nodes and experiences a combination of all separate transformations.

The dimensions of the user input have to be reduced with respect to the allowed degrees of freedom to consider the influence of the constraints. For objects that are not allowed to rotate, this is done by taking the displacement of the grasp center into account, rotation can be completely ignored. For an object that has all three degrees of freedom of translation this displacement can be directly applied to the object. For 2T-objects the displacement has to be projected onto the constraint plane. However, such objects do not occur in the car interior. For the existing 1T-restricted objects, correct object reaction can be easily achieved by projecting the motion vector onto the constraint vector. Therefore the travels $\overrightarrow{S}$ of all grasping pairs which are provided by the grasping heuristics, are averaged (Equation 3.13) and projected onto the constraint vector $\overrightarrow{CV}$ (Equation 3.14). This restricted translation $\overrightarrow{S_{restr}}$ is than — represented as the matrix M_s — multiplied with the matrix describing the last frame's pose of the object M_{t-1}. The resulting matrix M_t describes the location of the object as it was intended by the user (Equation 3.15).

$$\overrightarrow{S_{all}} = (\overrightarrow{S(gp_0)} + \overrightarrow{S(gp_1)} + ... + \overrightarrow{S(gp_n)})/n \tag{3.13}$$

Figure 3.11: Typical mounting constraints found in the car interior: 1R object mounted with a revolute joint (a), a 2R object, mounted with a saddle joint (b), a freely rotateable object, restricted by a ball and socket joint (c), a 1R object additionally movable with respect to a kinematic chain, defined by a 1R hinge joint and a 1T prismatic joint (d) and a 1R object, mounted with a hinge joint (d)

$$\begin{aligned} \overrightarrow{CV_{norm}} &= \overrightarrow{CV} / \left|\overrightarrow{CV}\right| \\ \overrightarrow{S_{restr}} &= \overrightarrow{CV_{norm}} \cdot (\overrightarrow{S_{all}} \bullet \overrightarrow{CV_{norm}}) \end{aligned} \tag{3.14}$$

$$M_t = M_{t-1} \cdot M_s \tag{3.15}$$

Rotational constraints are slightly more complicated to consider. It is necessary to understand how users interact with such objects to compute a plausible object reaction. There are basically two different ways of interaction with rotationally restricted objects. A lever-like rotation occurs if the grasp center has an offset to the rotation center of the object. The users rotate the object by moving the grasp center around the rotation center. One-handed turning of the steering wheel is an example of this kind of interaction.

Small controllers have to be handled differently. Here the users apply a rotation by turning the wrist of the hand. The grasp center then equals the rotation center of the object or lies close to the rotation axis. This rotation is called controller-like rotation. The derivation of user input from the translation of the grasp center would cause an incorrect object reaction in this case. Therefore it is crucial to use the rotation of the grasping pairs instead.

Figure 3.12: Rotation interaction: For big 1R objects both types of rotation — lever-like (left) and controller-like rotation (middle) — have to be provided, for 3R objects a combination of both rotation types results in plausible object motion (right)

For large objects, both types of rotation may have to be considered for the final object reaction. For example, the two-handed turning of the steering wheel usually is a combination of lever-like rotations of grasping pairs composed of phalanges of one hand and controller-like rotations of inter-hand grasping pairs (Figure 3.12). The same is true for 3R-restricted objects. Here, the object rotation is a combination of a lever-like rotation around the rotation center of the object and a controller-like rotation around the axis defined by the rotation center and the grasp center.

All possible interaction scenarios can be covered if the rotation constraints are considered at the grasping pair level. For 3R-restricted objects, both rotation types are calculated for each grasping pair. For the grasping pairs of 1R-restricted object, two separate cases have to be considered. A lever-like rotation is calculated if the grasping pairs barycenter is far away from the rotation axis and the rotation center. For grasping pairs with barycenters lying on or near the rotation axis or rotation center, the controller-like rotation is used. Allowing rotations of the other type would lead to incorrect results in each of the respective cases.

The lever-like rotation is derived from the translation of the grasping pair's barycenter around the rotation center. Therefore the vectors between the object's rotation center ($\overrightarrow{RC}$) and the grasping pair's barycenter is calculated for the current frame ($\overrightarrow{RB_t}$) and the previous frame ($\overrightarrow{RB_{t-1}}$) (Equation 3.16). Than the rotation axis ($\overrightarrow{RA}$) is calculated as the Cross Product of $\overrightarrow{RB_t}$ and $\overrightarrow{RB_{t-1}}$ (Equation 3.17). The rotation angle α is than calculated from the Dot Product of the vectors (Equation 3.18). For restricted objects ($\overrightarrow{RB_t}$) and ($\overrightarrow{RB_{t-1}}$) have to be projected into the constraint plane defined by the rotation center ($\overrightarrow{RC}$) and the constraint axis ($\overrightarrow{CA}$) of the object. Therefore the vectors are projected onto the constraint axis ($\overrightarrow{RBC_t}$ and $\overrightarrow{RBC_{t-1}}$ respectively). This projected vectors are subtracted from the original vectors to retrieve $\overrightarrow{RB_t}'$ and $\overrightarrow{RB_{t-1}}'$ (Equation 3.19). In this case the rotation axis equals the constraint axis.

$$\begin{aligned} \overrightarrow{RB_t} &= \overrightarrow{BC_t} - \overrightarrow{RC} \\ \overrightarrow{RB_{t-1}} &= \overrightarrow{BC_{t-1}} - \overrightarrow{RC} \end{aligned} \tag{3.16}$$

$$\overrightarrow{RA} = \overrightarrow{RB_{t-1}} \times \overrightarrow{RB_t} \tag{3.17}$$

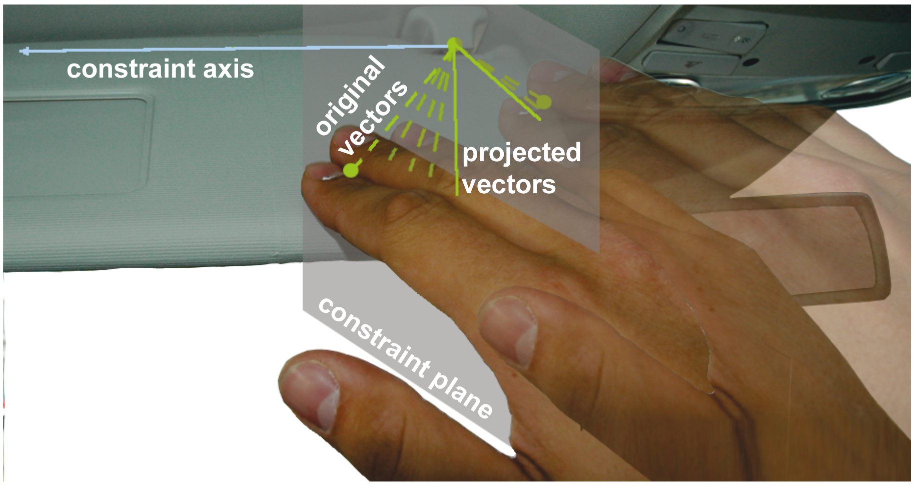

Figure 3.13: Restricted rotation: the original rotation vectors are projected onto the constraint plane before rotation angle calculation

$$\cos\alpha = \frac{\overrightarrow{RB_{t-1}} \bullet \overrightarrow{RB_t}}{|\overrightarrow{RB_{t-1}}| \cdot |\overrightarrow{RB_t}|} \tag{3.18}$$

$$\begin{aligned}
\overrightarrow{RBC_t} &= \overrightarrow{CA} \cdot \frac{\overrightarrow{CA} \bullet \overrightarrow{RBC_t}}{|\overrightarrow{CA}| \cdot |\overrightarrow{CA}|} \\
\overrightarrow{RBC_{t-1}} &= \overrightarrow{CA} \cdot \frac{\overrightarrow{CA} \bullet \overrightarrow{RBC_{t-1}}}{|\overrightarrow{CA}| \cdot |\overrightarrow{CA}|} \\
\overrightarrow{RB'_t} &= \overrightarrow{RB_t} - \overrightarrow{RBC_t} \\
\overrightarrow{RB'_{t-1}} &= \overrightarrow{RB_{t-1}} - \overrightarrow{RBC_{t-1}}
\end{aligned} \tag{3.19}$$

Controller-like rotations are calculated from the inter-frame difference of the line ($\overrightarrow{CL}$) connecting the two collision points ($\overrightarrow{CP}$) of the grasping pairs (Equation 3.20). This rotation is part of the rotational motion intention introduced by the grasping pair as explained in Section 3.1. Here, this rotation has to be reduced to a rotation around the rotation axis in the very same way as it was done for the lever-like rotation. This is done by projecting the two vectors into the plane perpendicular to the constraint axis ($\overrightarrow{CA}$) (Equation 3.21) (Figure 3.13). Other rotational influences of the grasping pair can be neglected since they cannot introduce a rotation around the axis. For 3R-restricted objects, the rotation axis defined by the line connecting the rotation center and the barycenter of the grasping pair is used as the constraint axis. In both cases the rotation can than be defined by the rotation or constraint axis ($\overrightarrow{CA}$) and the angle α which is calculated with the Dot Product of $\overrightarrow{CL_t}'$ and $\overrightarrow{CL_{t-1}}'$ (Equation 3.22).

$$\begin{aligned} \overrightarrow{CL_{t-1}} &= \overrightarrow{CP(f0)_{t-1}} - \overrightarrow{CP(f1)_{t-1}} \\ \overrightarrow{CL_t} &= \overrightarrow{CP(f0)_t} - \overrightarrow{CP(f1)_t} \end{aligned} \tag{3.20}$$

$$\begin{aligned} \overrightarrow{CLC_t} &= \overrightarrow{CA} \cdot \frac{\overrightarrow{CA} \bullet \overrightarrow{CL_t}}{|\overrightarrow{CA}| \cdot |\overrightarrow{CA}|} \\ \overrightarrow{CLC_{t-1}} &= \overrightarrow{CA} \cdot \frac{\overrightarrow{CA} \bullet \overrightarrow{CL_{t-1}}}{|\overrightarrow{CA}| \cdot |\overrightarrow{CA}|} \\ \overrightarrow{CL_t'} &= \overrightarrow{CL_t} - \overrightarrow{CLC_t} \\ \overrightarrow{CL_{t-1}'} &= \overrightarrow{CL_{t-1}} - \overrightarrow{CLC_{t-1}} \end{aligned} \tag{3.21}$$

$$\cos\alpha = \frac{\overrightarrow{CL_{t-1}'} \bullet \overrightarrow{CL_t'}}{|\overrightarrow{CL_{t-1}'}| \cdot |\overrightarrow{RB_t'}|} \tag{3.22}$$

The constraint of objects that are allowed to rotate in two dimensions (2R-restricted objects) is defined by kinematics that are build from two subsequent 1R joints. As mentioned before such kinematics can be represented by a scene graph hierarchy of two group nodes, each responsible for a 1-DOF-rotation around one axis. Each 1-DOF-rotation can be handled separately and the geometry node which is the child of such a hierarchy experiences the superposition of both rotations resulting in a 2R-restricted object behavior.

The individual rotation quaternion resulting from the input of each grasping pair are merged to compute the total rotational influence of a grasp by using the SLERP algorithm [Sho85]. For 3R-restricted objects both types of rotation have to be averaged per grasping pair. For 1R-restricted objects the average rotation simply is calculated as the mean of all angles calculated from the individual grasping pair rotations. With the constraint-based modification of each grasping pair's influence, the overall grasp rotation respects the objects constraints and covers all described types of rotations.

Functional Constraints

The previous section explained how the objects' motions are constrained by the way they are mounted to the car body. There are further constraints that are constructed by engineers to realize a certain function of the object. For example, a controller that can be rotated around one axis can be further restricted by lock positions at which the states of the controller are switched. Stop positions limit the object's movement beyond a certain point. Consequently, functional constraints cannot be found with freely movable objects.

Lock position constraints can have several variants. The objects can be allowed to perform continuous motions with states that are set at defined positions without any influence on the object's motion. Another possibility is that these positions are realized using a snap function.

Figure 3.14: Examples for functional constraints: unrestricted controller (left), freely rotatable controller with discrete rotation (middle) and a controller with discrete lock positions and stop positions, the controller cannot be rotated beyond (right)

Finally, the third option is, that the motion of the object is discrete, allowing the object to snap only into the lock positions. A stop position constraint is defined by a range the object is allowed to move within (Figure 3.14).

For the consideration of functional constraints, a ghost object is introduced, which moves as if no functional constraints apply. The offset of the current pose of the ghost object to a valid design position — a 3D vector for translational and an angle for rotational objects — can be used to check for violations of functional constraints and for their definition. The offset can be calculated for each frame or continuously tracked during interaction. The current offset of the ghost object is compared to the pre-defined positions in each consecutive frame. If the current offset equals a lock position, it is reached. For objects with a snap function, the constraint is extended to a snapping range, whereas for objects moving discretely, all snapping ranges have to adjoin each other.

As long as the ghost object is not influenced by constraint condition, its pose is transferred directly to the pose of the actually manipulated object. Under constraint condition, the object is set to the lock position. For stop constraints, the pose transfer is only performed within the allowed range.

Since the object stays in lock position until the users' input moves the object out of the Epsilon area, a difference between users' motion, finger motions and the actual object motion can only be observed if it is intended — in the snapping or stopping case. If the user is releasing an object when it is under constraint condition, the ghost object has to be set to the pose of the rendered representation to avoid inconsistencies.

Location Constraints

Location constraints are constraints that are caused by object-object interaction. An interactive object collides with another object — not necessarily an interactive one — and this collision works as a constraint for the object motion. These location constraints represent a limit for the pseudophysical metaphor. In handling object-object interaction, the use of a physical simulation is required. The processes resulting from object-object collisions are too complex to represent them on a pseudophysical basis.

However, a plausible reaction to this constraint type is possible for most of the situations dealt with here. The collision of a restricted interactive object with a static part of the car body usually causes the interactive object to stop. For example, the sun-shield colliding with the windshield stops its motion immediately, as well as the interior mirror colliding with the windshield. For these situations location constraints can be transformed into the functional constraint type of a stop position.

The plausible reaction of a freely movable object colliding with a static part cannot be realized using the proposed metaphor. However, such scenarios are needed for clearance analyses. In order to get around this, such collisions are visualized by highlighting the involved triangles. The interaction between two or more interactive objects is also not within the scope of the metaphor. Therefore the use of physical simulations is required.

3.3 Conclusion

This chapter described a novel pseudophysical interaction technique — direct finger-based interaction. This approach combines simple yet powerful grasping heuristics with plausible object behavior calculated geometrically as a reaction to the user input. Challenges of object selection by direct grasping were addressed by enhancing the rules of grasping by normal proxies and a pinch-sensitive finger tracking device. The behavior of car interior objects is highly influenced by constraints. A classification of these constraints and their implementation was described in this chapter.

With the proposed direct finger-based interaction technique it is possible to reliably and robustly manipulate virtual objects as the users are used to from reality. This realism is a prerequisite for the realization of functional assessments during the product development of the automotive industry. With a comprehensive user study the realism of this approach will be evaluated as well as its intuitiveness and robustness compared to conventional indirect controller-based interaction.

Chapter 4

Tactile Feedback at the Finger Tips for Improved Direct Interaction

A crucial requirement for fine-grained finger-based interaction is the precise transfer of the users hands and fingers into the virtual environment. As previously discussed, an optical finger tracking system (Figure 4.1) is used because it provides high accuracy at the finger tips. As long as video-based markerless finger tracking solutions are not capable of providing comparable precision and reliability, finger tracking hardware at the users' hands has to be tolerated. However, the existence of this hardware involves the chance to realize unobtrusive interaction support with additional hardware incorporated into the finger tracking system.

This chapter discusses the potential of an integration of tactile feedback at the finger tips. Therefore an approach using Shape Memory Alloys is introduced that realizes pressure-based grasping feedback and that was publicized in [SMF07]. The advantage of this approach is that it is very similar to the sensation one has during grasping in reality where a pressure can be felt when the object comes into contact with the fingers. The disadvantage is, that a pressure sensation is rapidly masked by the human sensory system. On the contrary vibrational sensations can be percepted continuously but they often appear to be bothersome for the users. This chapter describes the development of the innovative pressure-based tactile feedback with the advantages and challenges of this approach compared to conventional vibro-tactile feedback. Additionally, a vibro-tactile feedback device is described that realizes the tactile sensation at the same location as the pressure-based feedback — at the users' finger tips.

4.1 Human Haptic Sensoric System

This section introduces a short and rough overview over the human sensoric system to get an idea of how it processes tactile information provided at the finger tips and to identify the challenges for tactile feedback devices. First of all, the human skin is responsible for protecting the body from several environmental influences, such as injuries, temperature, microorganisms and harmful substances [CMH05]. What is more interesting for the development of tactile displays is that it also works as an external sensory organ retrieving tactile and thermo-receptive information and the sense of pain. The focus of this work is on tactile devices mounted to the hands and fingers, which is why this chapter deals with the hairless skin as it is found on human hands.

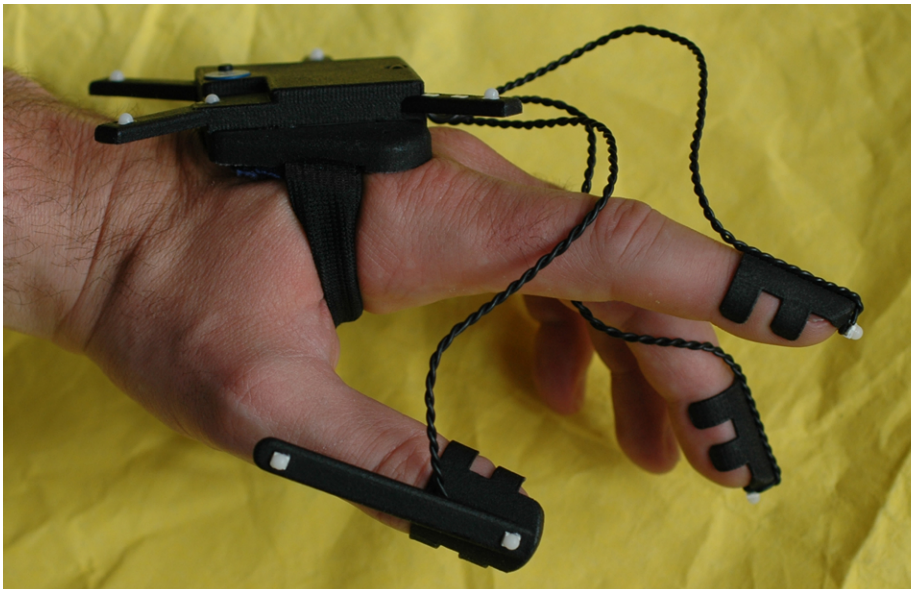

Figure 4.1: The optical finger tracking hardware by the A.R.T. GmbH provides the estimation of the hand configuration with the three most important fingers — thumb, index and middle finger — with the help of one active 6D target at the palm and four sequentially flashing 3D markers at the finger tips. This system also is available with 3D markers and finger bending estimation for all five fingers.

The outer layers of the human skin — the epidermis and the dermis — are interspersed with four different neuronal sub-systems for the reception of tactile sensations. Each of these mechanical-receptive fibers is sensitive to a certain kind of stimulus due to the properties of the receptor cells they consist of. A broader comparison of the capabilities of each receptor type with specifications of tactile displays can be found in [BHB^{+}95, CMH05]. Figure 4.2 shows a closeup of the skin with the most important cell types. The sensor types can be classified due to the rapidness of adaption to tactile stimuli. Two of the sensoric systems are slowly adapting (Merkel and Ruffini cells). The others are rapidly adapting (Meissner and Pacini corpuscle), which means that they only fire when a change of stimulus is detected.

Pressure on the skin and its indentation are detected by two of the sensor types. On the one hand, Merkel cells detect pressure on the skin or its indentation with high resolution. They slowly adapt to continuous pressure at low frequencies. They have very small receptive fields making it possible to differentiate stimuli that are located very close to each other. The receptive fields with Meissner corpuscle also have a high density on the fingertips and therefore a high spatial resolution. They sense very low frequencies below $40Hz$, which is often referred to as fluttering. They do adapt slowly to such a subsequent vibration-stimulus, but they adapt rapidly to a subsequent skin indentation. That is why pressure sensations can be differentiated with a very low two-point threshold, usually around $2.5mm$, but a subsequent pressure stimulus is quickly masked by the sensoric system and is hard to perceive.

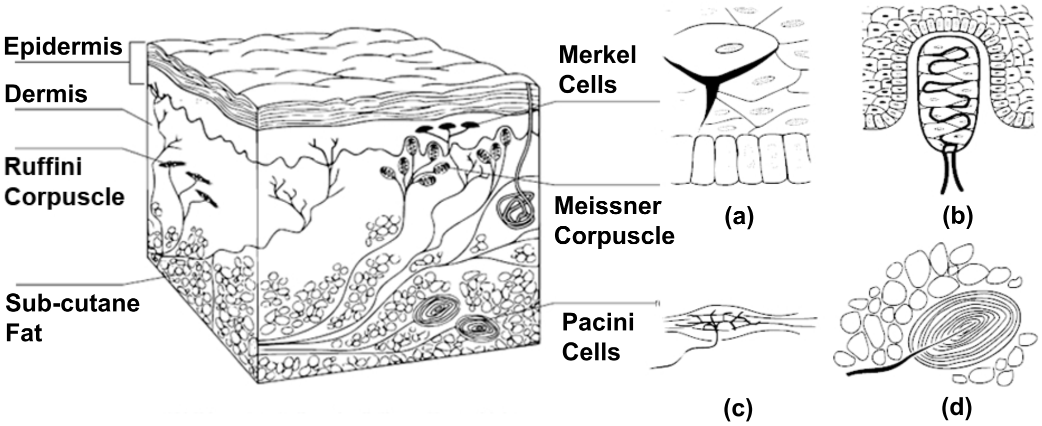

Figure 4.2: Structure of the outer skin layers with their most important elements (left) and a closeup of the four sensor types responsible for tactile sensations — Merkel cells (a), Meissner corpuscle (b), Ruffini corpuscle (c) and Pacini cells (d) [Gol03].

Vibration is detected by three of the cell types, sorted by the frequency of the vibration. As previously mentioned, fluttering with low frequencies is detected by Meissner corpuscle. Pacinian cells react on vibration with a frequency of more than $200Hz$. These cells are not as widespread as the Meissner corpuscle and their spatial resolution is not as high. Ruffini cells are responsible for horizontal displacement at low frequencies. Other work refers to slightly different parameters of the four cell types [JNL04, KH95, Gol03]. However, they agree that high frequency vibro-tactile stimuli are perceived by receptive fields that have a large diameter. The large size of these fields result in a JND (Just Noticeable Distance) that makes it impossible to differentiate more than one actuator at the finger tips [KH95].

4.2 Shape Memory Alloys — Characteristics and Applications

The requirement to have unobtrusive and light-weight actuators that can be incorporated into the finger tracking device led to the idea to use NiTiNol wires for the pressure generation. These wires are made from a shape memory alloy (*SMA*) of Nickel and Titanium. Shape memory alloys are characterized by the ability to alter the shape due to temperature changes. This special character provides the possibility to realize light-weight and unobtrusive actuators only consisting of the wires and a specialized electric control system.

Shape memory alloys have two distinguishable crystalline structures — Martensite, the low temperature state and Austenite, the high temperature state. By heating the material it can be forced to turn to Austenite, while when cooling down, it will return to Martensite. The thresholds of transformation depend on the mixing ratio of the alloy elements. The conversion is characterized by a hysteresis which means that the thresholds depend on the direction of the

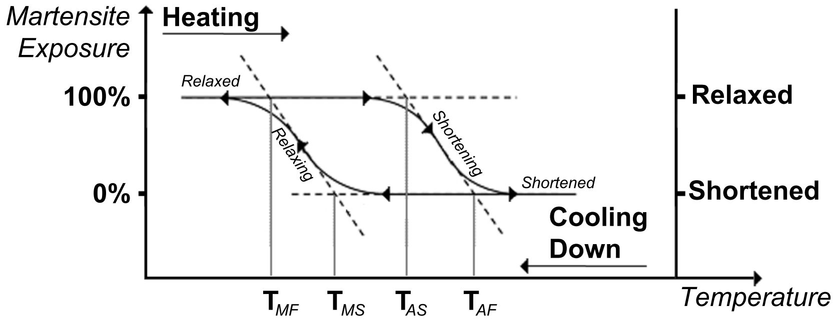

Figure 4.3: Characteristic hysteresis curve of NiTiNol. It shows that the shortening and relaxation processes happen at different temperatures.

transformation (Figure 4.3). When the increasing temperature reaches the threshold Austenite Start (T_{AS}), the material starts to transform to Austenite until it reaches the Austenite state with a temperature, referred to as Austenite Finish (T_{AF}). The cooling down Austenite material starts its transformation back to the Martensite state with the threshold Martensite Start (T_{MS}) and finishes this transformation when the temperature Martensite Finish (T_{MF}) is reached. Hysteresis is characterized by a disparity of T_{AS} and T_{MF} and T_{AF} and T_{MS} respectively. This special attribute of shape memory alloys makes it impossible to simply derive the material state from temperature. That is why it is difficult to build a simple control loop when components made from SMA are used.

SMA components can be forced to have a certain shape in each state by training them. As a simple example the shapes can feature different lengths of a wire for the two states. By forcing them into the desired shape while quickly heating with high temperatures they can be trained to take on more sophisticated shapes.

There is a trade-off between elongation length and durability of the material. If only one contraction is needed it can provide length changes of up to eight percent. For a two-way effect elongations of up to five percent are reported but this is paired with a reduced lifetime of tens of thousands of cycles or less. With durability as the main focus, only around two percent of length change should be trained [JNL04]. Another property of SMAs is pseudo-elasticity, meaning that the object can be deformed and the material returns immediately to the former shape. Here Austenite is built mechanically leading to a strain of up to eight percent which is more flexibility than any other metal can provide.

Shape memory alloys are well known and have been used in several research projects and commercial products for years. All three properties, the one-way effect, pseudo-elasticity and the two-way effect, are useful for certain requirements. Although there are applications for the one-way effect (used for the sealing of pipe connections) and pseudo-elasticity (used with unbreakable glasses frames), this work focuses on approaches using the two-way effect since this is necessary for a tactile feedback device.

Besides applications in commercial products, such as temperature sensors, there are few other

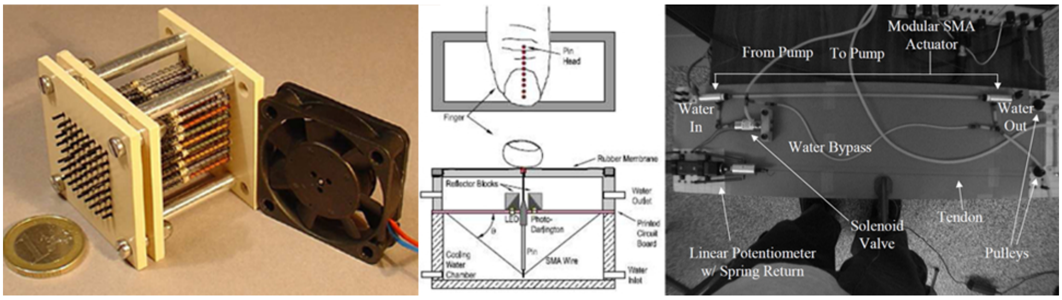

Figure 4.4: A small tactile display with pins driven by SMA wires [VPHS05] (left), a single SMA-driven pin [WPFH98] (middle), the complex apparatus necessary for one liquid cooled SMA wire [MA03] (right).

approaches for providing tactile feedback based on shape memory alloys. Arrays of pins used as tactile displays are presented in [WPFH98, VPHS05] (Figure 4.4). These devices can be attached to input devices to provide additional information. Jones and Nakamura mounted an actuator in a vest which stimulated the users skin by pins driven by SMA wires [JNL04].

Shape memory alloys have one particular disadvantage for the use in tactile feedback devices. While transformation to Austenite usually occurs quite quickly, returning to Martensite is a slow process since it depends on cooling-down of the material. Thin wires of NiTiNol can contract in less than $50ms$ but need time in the order of a second to relax. This delay can lead to misinformation since the state of the application differs from the state of the feedback device. Many ideas have been generated to overcome these problems. Since this is a cooling problem, suggestions were made to use cooling liquids [WPFH98, MA03] (Figure 4.4). An US patent [HL92] describes a method of producing specially conditioned fast twitching fibers driven by electromagnetic pulses. Alternatively actuators can be forced mechanically to the Martensite state, or SMAs having an intrinsic two-way effect speeding up the return path can be produced [Esc93]. These approaches have in common that they are often obtrusive and not very robust and some of them are not suitable for being used for a high number of shortening and relax cycles.

Although the material characteristics of Shape Memory Alloys imply a number of challenges, they have the potential to enable light-weight pressure-based tactile feedback at the fingertips.

4.3 Design of the Tactile Feedback Actuators

During direct finger-based interaction, the user's fingers come into contact with virtual objects and if the contact situation fulfills the requirements of a valid grasp, the object can be manipulated. Usually the virtual objects can be perceived by visual sense alone, making it difficult to judge the fingers' location with respect to the objects'. With an unobtrusive contact or grasping feedback at the fingertips, the reliability of the interaction might be supported considering that the users get additional information on the interplay between the virtual objects and their hands. The integration of lightweight tactile actuators into the thimbles of the optical finger tracking de-

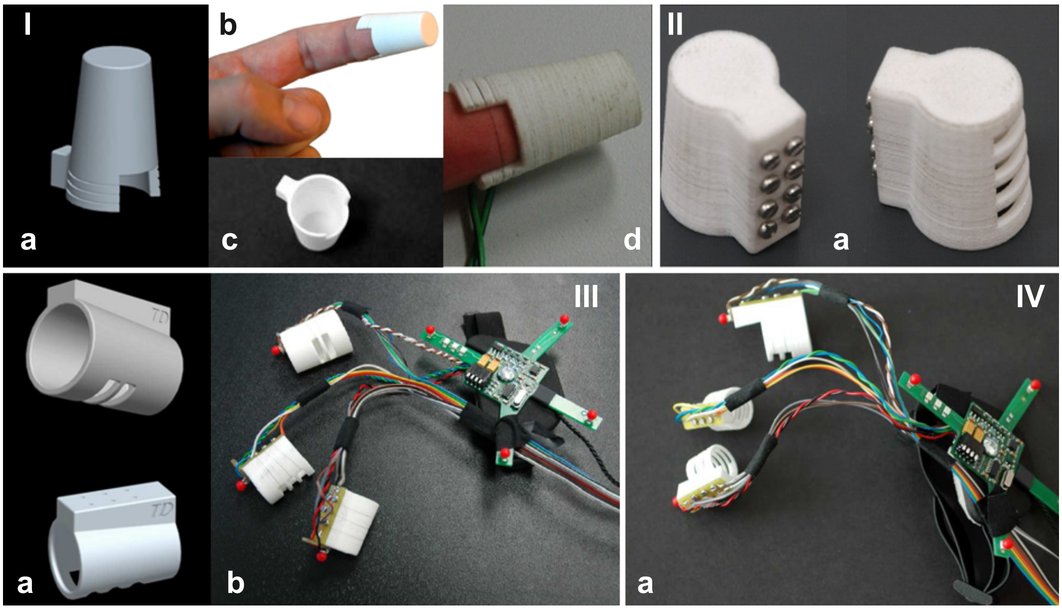

Figure 4.5: From concept to prototype: 1st study with one SMA-actuator (I) — the design (a) and its implementation (b-d) , tests with several actuators per thimble (IIa), 3-finger approach connected to the finger tracking system (III) and the final prototype using thimbles with reduced material (IV).

vice promises to be a sufficient way for the realization of such an unobtrusive feedback. The idea of this thesis is to use thin SMA wires wrapped around a thimble, which is open at the inner side of the fingertips allowing the wires to be directly in contact with the skin. The wire contraction creates a localized pressure sensation on the fingertip, which should be easy to perceive.

The thimbles were created in a selective laser sintering process using polyamide. Laser sintering is an additive process, which does not require special stilts. Thus the manufacturing process does not need to be considered during the design phase. Both, the mechanical and thermal stability of the material are very high. In contrast the relatively low resolution of 1mm and the roughness of the surface are disadvantageous, but are acceptable for prototypes. Four prototypes were developed with incrementally improved designs (Figure 4.5). The prototypes were limited to provide tactile feedback for the three main fingers of each hand – thumb, index and middle finger. These fingers are mainly involved in direct finger-based interaction, the ring and pinky finger often only provide grasping assistance. Consequently the optical finger tracking system, working as a basis here, usually is limited to these fingers to reduce the complexity and obtrusiveness of the system.

The first prototype was a single thimble housing one SMA wire. This design step was used for proof of concept, conveying if the user is able to feel the contraction of the wire. For the second design, the number of actuators per thimble device was increased. A clamped connection with $M2$ screws was used for the wires, so the maximum number of actuators per thimble mainly depended on the size of the screw heads. The screws' head diameter was $3.8mm$, so a spacing of

Figure 4.6: First evaluation of the final prototype with car interior interaction in a CAVE-environment.

$6mm$ was included between the wires. This decision allows for four actuators to be attached to the thimble. However, another design limitation arose from the device control board used, which was able to control only eight wires simultaneously. As a result, three actuators were integrated into the thimbles for the index and middle finger and two for the thumb. The third and the final prototype combined the tactile feedback actuators with the optical finger tracking prototype developed by A.R.T. GmbH [Web10a]. First evaluations in an immersive environment were performed with the third prototype leading to the final version with improved ventilation for the finger tips. Extraneous thimble material was reduced as much as possible without reducing the robustness of the finger actuator (Figure 4.5). The final prototype was used for first evaluations in a CAVE-environment (Figure 4.6)

A critical parameter is the fit of the thimbles on the users' finger tips. There is an ISO standard describing hand parameters — ISO 33402. However, some more parameters than specified are needed for building thimbles, namely the outer phalanges' lengths and circumferences are not defined by this standard. With averaging the hand parameters of eleven subjects — defined as typical users of the intended applications — common dimensions of the thimbles were chosen. For the middle and index finger an inner diameter of $17mm$ was used, narrowing to $11mm$ over a length of $28mm$. The thumb thimble has an inner diameter of $23mm$ narrowing to $18mm$ and its length is $35mm$. These thimble sizes fit most of the intended users. However, people with either rather big or quite small fingers would require the development of adjustable thimbles or thimbles of varying sizes would have to be provided.

Another challenge is the connection of the SMA wires to the power supply, which is commonly achieved through a proprietary soldering process. However, improper soldering can re-

Figure 4.7: First commercial product prototype of the pressure-based tactile feedback device as it was used within this work (left) with a close-up of the wires incorporated into the finger thimbles (right)

sult in losing the memory effect. Furthermore, easy replacement of the wires is required. That is why the wires were clamped with screws at the back of the thimble.

As previously mentioned, only small length changes can be realized if a long life time of the actuators is priority. One of the first tests aimed at the recognizability of these small wire contractions. It revealed that a $50mm$ long wire wrapped around the first thimble prototype shortens about $1.5 - 2.5mm$ when a current was applied. Users were able to notice the effect without problems even though this creates only a very small impression on the skin.

Furthermore, some experiments were performed with individually controlling the three SMA wires of a single thimble. Thereby it was noticed that the impressions of the different wires can be distinguished very well. Users can even feel patterns displayed by activating the actuators sequentially. However due to the design and control of the prototype, the number of actuators is limited. A more elaborate fastening mechanism and a more advanced controller design would allow for the adding of more actuators to each thimble. The two-point threshold of the fingertips is around $2.5mm$. So actuator arrays with interspaces larger than this discrimination limit will generate discretely perceivable signals.

The final thimbles were integrated into a prototype of the optical finger tracking system developed by A.R.T. GmbH. This system consists of a hand target with four active markers and one active marker for each finger. The finger markers are usually attached to plastics thimbles. Instead they were glued to the prototype thimbles as shown in Figure 4.5. This prototype only supported three tracked fingers per hand. The incremental design process led to a prototype consisting of a small backbone for holding the screws and thin rims around the finger. There are three actuators for the middle and index finger and two actuators for the thumb. This final design showed a good compromise between stability and wearing comfort. The thimbles are lightweight and easy to put on. The ring and pinky fingers are not yet supported, but they are less relevant for most of the intended interactions. In the meantime this approach was further developed by the A.R.T. GmbH. The basic principals have not be changed but the design of the housings have been refined to allow for commercial distribution (Figure 4.7).

4.4 Device Control

Several experiments with a set of different alloy types and diameters were conducted to select the appropriate wires and identify the control parameters for the electrical design. To achieve a rapid transformation the actuators have to be as thin as possible, of course limited by the tensile strength of the material. Furthermore, thin wires require less time to cool down, which is important for a short relaxation process. Tests showed that $80\mu m$ wires have sufficient tensile strength and showed the best dynamic performance. The shape memory alloy is directly heated by electrical current using a pulse width modulation (*PWM*) signal to initiate the shortening process. The skin was utilized as a heat sink for the wires. It is important to note that there is no risk for the users to contract burns if the actuators are thin enough and the transformation temperature is reached within a short time — for the $80\mu m$ wire in less than $50ms$.

As a result of the experiments, a pre-trained SMA type M was used that was purchased from the Memory-Metalle GmbH, Germany, that is now part of the Memry Corp. [Web11f]. It had a diameter of $80\mu m$ and a length of approximately $50mm$ per actuator. This material has a T_{AF} of around $65°C$. Due to the characteristics of the sensory system, it has to be expected that an ongoing impression of the fingers' skin — reporting an ongoing collision for example — can only be recognized by one of the responsible sensor cell types. The impression generated by the SMA actuators is a stimulus which is not strong enough to be continuously perceived. However, applying a low duty-cycle of the PWM signal results in a perceivable pulsation of the material. This vibration is used to provide a second tactile pattern. The material is transformed and contracted using a high duty-cycle for a short period of time. Afterwards a low duty-cycle is applied producing an ongoing stimulus. During the low duty-cycle the wires already cool down and relax, which remains unnoticed by the users. Stopping the low duty-cycle and thus the pulsation at the end of a collision with a virtual object is immediately recognized. This way the slow response time for the relaxing process of the SMA wire is obscured by the pulsation signal. However, this pulsation cannot be perceived comparably well by all users. Furthermore, its strength changes over time, probably due to material fatigue. Thus, this pulsation of the shape memory alloy cannot be used as a reliable tactile pattern for ongoing stimuli.

Another important fact to consider is power consumption. The power transferred by a PWM signal is calculated as explained in Equation 4.1. The wires are triggered time-multiplexed, so the maximum duty-cycle per actuator is one divided by the number of actuators — eight wires in this case. Experiments revealed that the $80\mu m$ wires with a resistance of 11Ω can be excited within $40ms$ running at $5V$ and using an 80% duty-cycle. This results in power requirements of $1.81W$ following equation 4.1. A voltage of at least $12.6V$ has to be applied for achieving the same power while multiplexing with only $1/8th$ maximum duty-cycle (Equation 4.2). The control board was designed to power the eight wires with $13.8V$.

$$\begin{aligned} P &= \frac{U_{on}^2}{R} \cdot \frac{t_{on}}{t_{on}+t_{off}} + \frac{U_{off}^2}{R} \cdot \frac{t_{off}}{t_{on}+t_{off}} \\ & if\ U_{off}\ is\ 0 \\ P &= \frac{U_{on}^2}{R} \cdot \frac{t_{on}}{t_{on}+t_{off}} \end{aligned} \tag{4.1}$$

$$\begin{aligned} U_{new} &= \sqrt{1.81W \cdot 11\Omega \cdot 8} \\ U_{new} &= 12.6V \end{aligned} \tag{4.2}$$

For communication with the VR application, a micro-controller based circuit board was developed that allows to drive the time-multiplexed wires with a PWM signal. Therefore an ATMega8 micro-controller was used running at $8MHz$. The ATMega8 from Atmel Corp. is a RISC processor with $8kByte$ flash program memory, $1kByte$ SRAM, 512 byte EEPROM, 6 or 8 channel 10 bit A/D-converters, providing 8 MIPS at $8MHz$. It is programmable using the assembler instruction set or the GCC tool chain. The firmware for this application was developed in C using the avr-libc version $1.4.4$. A command line application waits for commands sent via UDP and sends commands to the microcontroller via radio.

The micro-controller generates 3 signals:

- A PWM signal for each actuator channel
- A signal to switch the channels
- A reset signal for the channel counter on the printed circuit board (PCB)

A 4017 decade counter is used to switch between the eight channels. The switching clock for this circuit is provided by a timer 0 overflow interrupt of the micro-controller. After eight switching cycles a reset signal is sent. A separate PWM duty-cycle is applied during the timeslot for each channel. It is possible to extend the board design to control more wires (e.g. to support 15 wires, 3 for each finger tip), but the voltage level has to be adapted as was indicated above. Alternatively a micro controller with more PWM channels could be used.

4.5 Vibration-based Tactile Feedback Device

A different kind of tactile feedback is realized with the vibration-based tactile feedback device (Figure 4.8). Its development was motivated by the main drawback of the pressure-based feedback system. As described in a previous section, it is not possible to give the user ongoing pressure feedback during the whole interaction process. For the vibration-based system, voice coil motors — as they are used for the vibrational alarm of cell phones — are incorporated into the finger thimbles instead of SMA wires. The overall system setup is comparable to the pressure-based system.

Figure 4.8: Vibration-based feedback device (left) with a close up of the voice coil motor incorporated into the finger thimbles (right)

This vibration is handled by different sensor types and can be continuously felt by the users. Thus this system is capable of providing feedback on the appearance, duration and disappearance of grasps. On the other hand, the sensoric system responsible for vibration detection is relatively large and the sensation easily couples into the finger bones. That is why the location of the feedback cannot be discriminated perfectly when vibration sensations are used. Moreover, this kind of feedback is potentially more obtrusive than the pressure sensation.

However, as was evident with the pressure-based system, the feedback again is located where the interaction takes place — at the finger tips. It can be expected that this is an advantage over existing systems, such as the commercially available CyberTouch™ for example, that in history were considered to be uncomfortable by VR users. It will be interesting to see if the tactile feedback will help the users during interaction and which approach will be preferred.

4.6 Integration into the Software Environment

Both devices proposed for tactile feedback have been integrated into the VR system. This particular VR-system has a device manager that handles the initialization and control of input devices. Considering that the tactile feedback systems are in- and output devices, it seems obvious to integrate them as devices in the device manager. However, the extension capabilities of this component are reduced and it only can handle input devices. So the device driver for the tactile feedback had to be realized as an extension to the interaction manager.

The modules responsible for device integration were developed as dynamically shared objects (DSO). Init functions are used to set parameters of the module and to open communication ports. The control of the feedback devices is realized in a callback function, that is called when a valid grasp is detected by the grasping heuristics. Depending on which finger is involved in the grasp, a two byte command sequence is sent to the micro controller. The first byte contains the command itself:

- Contact start

- Contact end
- Set contact-start PWM duty-cycle
- Set contact-hold PWM duty-cycle
- Set contact-start duration

The second byte supplies the following values, if required:

- Information on which wire has to start or stop
- Values for PWM duty-cycle and duration

Since these values are 8bit wide, a maximum of 255 wires can be addressed, and the PWM duty-cycle is also adjustable in the range of 0 and 255.

4.7 Conclusion

In summary, the fact that the NiTiNol wires possess a shape memory effect can be used to realize light-weight tactile actuators that can be incorporated into the finger thimbles of an optical finger tracking system. Minimal hardware is necessary to generate the pressure feedback and it can be expected that the users can be supported with direct finger-based manipulation of virtual objects without disturbing them during interaction. Within this thesis a prototype of such a tactile feedback system was developed that is now a commercially available device. The pressure-based feedback realized with this device has the advantage that it can be easily perceived by the users and that it is very close to what the users feel when they touch real objects. The disadvantage is that ongoing feedback cannot be realized since the human haptic sensoric system rapidly adapts to this stimulus. This is not the case with vibro-tactile stimuli as they are provided by the second device introduced in this chapter. However, the vibration generated by this device can be bothersome and might disturb users during interaction. The amount of user support provided by tactile feedback in general and the user preferences on the kind of feedback will be further evaluated.

Chapter 5

Evaluation

The previous chapters described the direct finger-based interaction technique that enables the functional validation of car interiors by providing a realistic yet reliable manipulation of virtual objects. The interaction technique has to be carefully evaluated in order to prove this special characteristic. This evaluation is done with three studies that are described in this chapter. The first study was publicized together with direct finger-based interaction in [MF10]. The second study, a comprehensive user study, was described in [MF11a]. The third study was part of the publication [MF11b].

5.1 Usability Analysis

As an initial study, the usability and realism of the pure interaction technique was tested without considering the proposed enhancements of the finger tracking hardware. Therefore an expert review was performed, since there is no quantitative measure for these issues. Expert reviews are a common usability evaluation technique in software development, usually revealing 75% of all usability problems with the help of only five experts [PRS02].

5.1.1 Method

Participants

The expert review was performed with six experts for automotive virtual reality applications. An even number of experts was chosen to allow for paired reviews that enable productive discussions. The experts were recruited from various departments of the company. They were ergonomists, simulation experts and automotive VR system developers.

Apparatus

The expert review was conducted in a three-sided CAVE. This CAVE is driven by a PC-cluster generating images for 10 full-HD projectors. A pair of two projectors provides a passive stereo image separated by the INFITEC technology. Two pairs of projectors are vertically arranged for

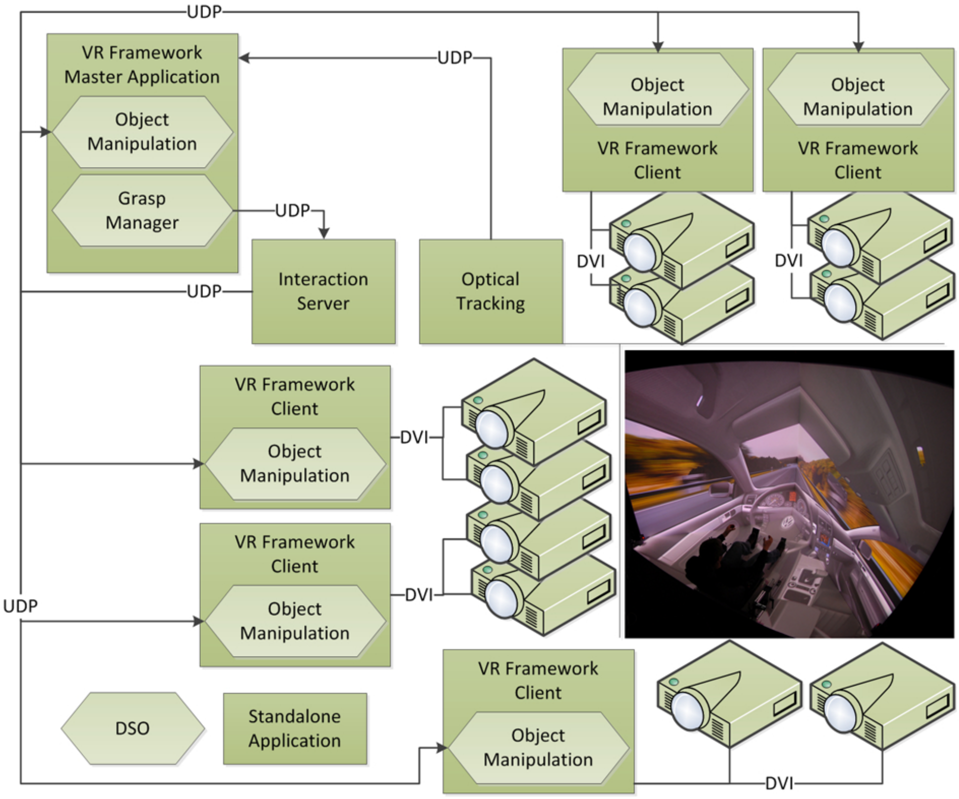

Figure 5.1: The interaction framework architecture.

each of the side walls with horizontal edge blending, resulting in a total resolution of 1920x1920. The floor image is diagonally projected by one projector-pair from above.

Each projector pair is connected to a workstation responsible for rendering the respective image. Therefore an instance of the VR framework is running on each of the rendering clients holding a copy of the virtual environment that is accessible from a common network resource. The actual application, managing in- and output devices, the interaction with the virtual environment and all other issues, is running on a master workstation. This master application distributes the inter-frame differences of the camera and the input devices to the clients via network. This cluster support is originally provided by the VR framework and is used as a basis for the developments presented within this work.

The interaction framework that was conceptually described in the preceding chapters was realized based on the VR system Virtual Design 2. This particular system allows for extensions by loading Dynamic Shared Modules (DSO) that provide well defined interfaces (see Section 2.5). For the realization of all aspects of direct finger-based interaction, a couple of modules and softwares were developed that interact with each other (Figure 5.1). The grasp manager collects information from the system-inherent collision detection to detect and manage grasp intentions of the users. It gathers the user input during manipulation and provides this information via

Figure 5.2: Prototype application with a user manipulating the interior mirror in the CAVE

UDP to an object reaction server running independently from the VR system. This server is provided with knowledge about the constraints of each manipulatable object via a config file and calculates the pseudophysical object reaction on user input.

This client server design was chosen for providing cluster support which is necessary for the desired display systems. Since the original VR system is not capable of managing object manipulations within a cluster framework, the resulting object matrices have to be distributed to all rendering clients (and the master) to guarantee consistent object states on all partial screens. Because of the loose coupling and the missing synchronization, small differences and delays happen to occur between the screens. Since the interaction server always provides the latest object matrix and due to the relatively high update rate, the differences are small and negotiable. Moreover, these small differences have to be preferred over an overall system latency that would have been caused by a more sophisticated synchronization method. Each of the client-applications and the master-application run a DSO that is responsible for applying the latest manipulated matrices to the respective virtual objects.

Precise finger tracking is realized through the wireless finger tracking gloves from Advanced Realtime Tracking. Optical hand and head tracking is supported by the same camera system. As a test scenario, the interior of a Volkswagen Touareg was chosen consisting of approximately two million triangles (Figure 5.2). The whole application in the CAVE was running at around 12-15fps. Within this test scenario the following objects — each representing a group of objects with special characteristics — could be manipulated:

- Steering wheel
- Sun shields
- Glove box lid
- Interior mirror
- Light switch

- A freely movable soda bottle
- Driver's door

Design and Procedure

The expert review was performed in terms of semi-structured interviews with three pairs of experts. This design was chosen to allow for a free discussion among the interviewees. Each session started with a testing period of approximately 20 minutes for each of the experts. They were asked to interact with all available objects during this period. While they interacted with the virtual car interior, the experts already discussed the prototype. After each subject experimented with the interaction metaphors, they were interviewed together for another 20 minutes. As a guide throughout the interview, a number of questions were prepared and casually asked during the discussion. These questions addressed the main challenges for the usability of the interaction metaphor:

- Are you able to grasp the virtual objects easily?
- Are you able to release the virtual objects easily?
- Do you identify differences concerning the virtual objects?
- Do the objects behave as you expected?
- Are you able to judge the position of your hands and fingers with respect to the virtual objects?
- Are you able to evaluate the accessibility of the virtual objects?
- For which aspects of virtual car validation do you consider this kind of interaction to be appropriate?
- For which particular aspects could this metaphor be used in contrast to an indirect controller-based method?
- Which improvement suggestions do you encounter and which application areas would open up as a result?

5.1.2 Results and Discussion

Judging the actual grasping, the experts stated that they were able to select objects very intuitively and robustly with the help of the grasping heuristics. Problems occurred with smaller objects due to occlusion. Interestingly, it turned out that grasping the objects with a tip pinch worked best for the users, considering it provides least occlusion and the most precise way of interaction.

Concerning the release of an object, the experts considered it mandatory to carefully balance the stop condition parameters. Sometimes objects appeared to be sticky and sometimes objects tended to follow the finger motion even though the users already opened their hand with the intention to end the grasp. The threshold of the first stop condition (see Section 3.1.3) was responsible for both of these effects. Given a generous threshold, the grasping heuristics are not able to rapidly detect the users' grasping intentions. However, a small threshold can result in unstable grasp continuity due to jitter and imprecise finger motion. In both cases unacceptable releasing behavior is induced having the potential to quickly annoy the users. This trade-off could be solved by a calibration of this parameter or a careful adjustment per object. However, experts became quickly proficient with the technique and precisely released objects by dropping them carefully. Furthermore, the second stop condition was appreciated as a definite interaction stop in cases where objects appeared to be sticky.

The motion of the objects as a reaction to finger collisions was judged to be very plausible and realistic. In particular, the precise response to very fine grained motion of the finger tips fascinated the experts and was highly recommended. Multi-hand interaction was successfully tested by passing the bottle from one hand to the other. The same would be possible with our method for two or more users in a multi-user stereo setup.

The push interaction was also judged to be very intuitive since object reaction immediately follows a collision. However, the experts complained about the missing mass and inertia of the objects. At first sight it was unrealistic for them that objects did not move according to their size and mass when touched by the user. Also the missing mass and inertia of virtual objects that are grasped can appear to be unrealistic. However, unless appropriate haptic feedback devices are available, it is not possible to display forces coming from the virtual environment to the user. Without this haptic rendering, pseudophysical feedback of the objects taking their dynamic properties into account would appear to be unrealistic due to the direct character of the interaction metaphor. However, it could be interesting to think about pseudophysical force rendering as it was shown for indirect manipulation by Lécuyer et al. [LCR$^+$00].

With direct finger-based interaction, the experts were able to judge the accessibility of objects. In contrast to the indirect method they were used to — involving an input device such as the Flystick — they were able to select the objects directly with their hands. Moreover, they found the direct method to be more intuitive, since no button assignment needed to be learned. Due to the direct interaction involving the hands and arms of the user, the experts identified a potential range of new applications for these techniques. This range included assembly simulations which need to consider the clearance of the human hand — especially concerning car maintenance — to the assessment of ergonomics issues in the car interior. Furthermore, the experts stated that this interaction would enrich immersive design and concept reviews due to the increased interactivity and immersion, enabling a more realistic experience of the virtual model. However, the direct object selection and release method was stated to be less precise than pressing and releasing the button of an input device.

An interesting question raised was, if the experts would prefer to use a virtual representation of their own hands during interaction. Despite being occluded by the real hands, the virtual hands can be seen by the users. The reason for this is that a simple hand model was used, having average finger diameters and a solid palm. This visual feedback was preferred by half of the experts, giving them the opportunity to judge their virtual fingers' relation to the objects. The other half preferred to interact without a hand representation. They appreciated that they were able to concentrate on the virtual object without being distracted by the virtual hand representation. They further noticed that because they were not tempted by the virtual hand model to look at their real hands, they had fewer problems with the focus and convergence mismatch. Since the expert review did not reveal a clear preference, a further investigation of the influence of the virtual hand representation on grasping robustness became necessary.

Finally, the experts remarked that for the assessment of object clearance in very complex maintenance and assembly scenarios, object-object interactions through collisions are required. These object-object interactions cannot be realized with the pseudophysical approach presented here. These complex interaction processes cannot be realized without simulating the physical interplay between the users and the objects.

5.1.3 Conclusion

The expert review was performed for evaluating the usability of direct finger-based interaction within a test application. It proved that the experts were able to reliably and intuitively interact with the virtual objects. The high realism and intuitiveness of this method were explicitly emphasized by the subjects. The applicability for virtual functional assessments of the car interior were explicitly emphasized by the users as well as the potential to enrich assembly simulations and concept reviews. Difficulties occurred with reliably grasping small shapes. These objects, such as the light switch, turned out to be challenging for direct interaction, as was expected. These challenges were the target of the hardware enhancements which will be evaluated in the next chapter. An interesting issue was that the experts were ambivalent concerning the benefit of a virtual hand representation for interaction. This issue will also be further evaluated in the next chapter.

5.2 User Study

The expert review that was conducted as a pilot study, conveyed a first impression of the usability of the direct finger-based interaction as it was realized within this work. Moreover the potential and limitations of this approach as well as the range of possible applications could be discussed with the experts. On one hand, the experts appreciated the intuitiveness and realism of this technique, but on the other hand, they considered conventional controller-based interaction to be more robust and reliable. For a more thorough understanding of the respective advantages of both approaches it was inevitable to quantify the realism and the ease-of-use of the techniques.

Within this work a number of in- and output devices were suggested to support the users during direct finger-based interaction and to overcome some of the limitations of such direct interaction metaphors. For the verification of these assumptions, the amount of user support provided by the proposed devices had to be quantified.

At the Volkswagen Group, mainly projection-based immersive environments are used for functional assessments of car concepts because of their good user acceptance resulting from the large field of view and decent comfort. Nevertheless, there are also applications which require the use of HMDs (e.g. the evaluation of the customers' experience utilizing Virtual Seating Bucks). Literature indicates that the choice of the display system could have an impact on the user performance and the usability of the interaction metaphor.

Motivated by the described challenges and questions, an extensive user study was designed that addressed the following influences on the interaction metaphor:

- Comparison of direct finger-based interaction with an indirect metaphor using a controller
- Influence of the choice of display system on the interaction experience
- Amount of user support by enhancements of the finger tracking hardware — pinch sensitivity and tactile feedback

Concerning the choice of measures for the user study it was considered to be important to cover all dimensions of the interaction experience. That is why quantitative and qualitative measures — task completion times and subjective user judgments — were used as well as a survey for direct user preferences.

5.2.1 Method

Participants

Twelve subjects participated in the assessment, eleven of which were male and one female. Their ages ranged from 22 to 39. Two of the subjects were left-handed and the rest right-handed. All participants had unimpaired or at least corrected sight. The group involved VR system experts as well as experienced VR users and undergraduate students considered to be novice users. This spectrum is representative of the typical mixture of user background for industrial virtual reality applications.

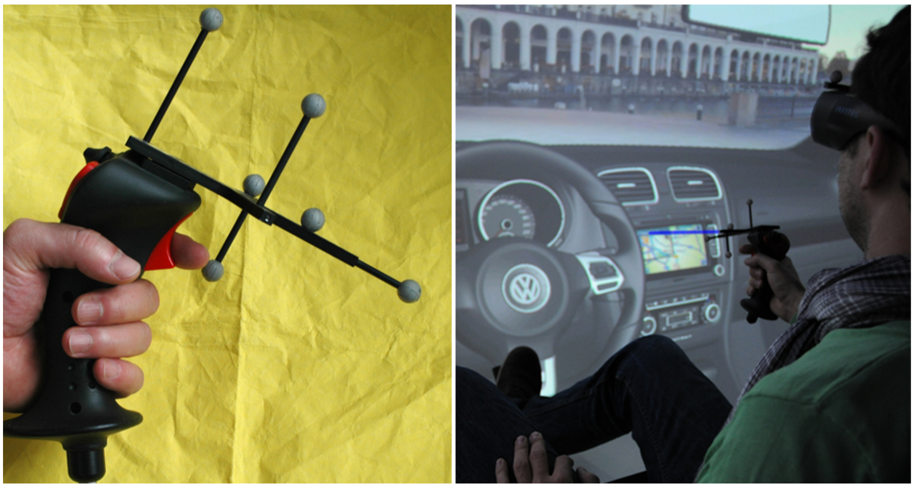

Figure 5.3: The controller used for indirect interaction — a Flystick incorporating a 6dof-tracked handle and several buttons (left) and a typical controller-based interaction scenario using ray-based picking (right).

Apparatus

Display Systems For the user study both display types were used which are usually employed in the targeted applications. On the one hand, the three-sided high resolution CAVE was used that was previously described in Section 5.1.1. It provides a wide field of view, almost covering the whole natural human field of view. The satisfying visual quality and user comfort lead to a high user acceptance of this system. For direct interaction metaphors, projection-based VR displays are challenging. Mainly occlusion effects and focus shifts between the real hands and the projection screens make it hard for some users to judge the location of the user's hands with respect to the virtual objects. Thus, reliable hand-object collisions needed for robust finger-based interaction are not always easy to achieve.

On the other hand, Virtual Seating Bucks (cf. Section 1.2) complement minimal hardware mockups usually consisting of a steering wheel and a driver seat with virtual car data presented by a Head Mounted Display. These displays are capable of fully immersing the user and exclude any real environment as well as the user's own body from the perceived image. Consequently, the judgment of hand-object relations are entirely based on the virtual hand representation and challenges of direct interaction, such as occlusion effects, are avoided. The drawback of these displays is their limited field of view and a lack of comfort. The HMD used in this study is a carefully set up Rockwell Collins SR80 providing 80° diagonal FOV with 100% overlap and SXGA resolution. This particular display is quite comfortable with a low and balanced weight. Virtual hands were not hindered from penetrating virtual objects. This might appear unnatural and thus disturb the users, but the users did not complain about this fact and accept this kind of interaction in HMD applications.

Controller-based Indirect Interaction Utilizing a Flystick As a reference interaction method, an indirect controller-based metaphor was implemented. Therefore a Flystick (Figure 5.3) was integrated into the application, an input device commonly used for immersive industrial applications (e.g. for assembly simulations). It is a tracked joystick-like handle with several buttons that transmit their state via radio. A virtual cursor — a 3D-model of a ray — attached to the physical device is used to interact with the scene. "Grasping" is realized by ray-based picking requiring the user to pierce the desired object with the cursor and to press the "Fire"-button. The objects' reaction on indirect user input is the same as it is for direct input. Therefore two reference points were attached to the picking geometry with an offset of $5cm$ to each other. These reference points were used as if they would be finger segments that define a grasping pair for the selected object. The described functionality was integrated into the grasp manager module of the implementation.

Controller-based indirect interaction can be very robust and easy to use. The button-based selection provides very clear feedback on interaction start, stop and status. Even novice users understand this interaction metaphor immediately. Consequently, ray-based picking with a Flystick is the commonly used interaction technique for industrial applications. On the other hand, the metaphor is very abstract and does not provide the realism needed for assessments concerning accessibility and usability in car interiors. The user study is expected to reveal the individual advantages and disadvantages of each interaction approach.

Pinch-sensitive Grasping Furthermore the pinch-sensitive finger tracking device introduced in Section 3.1.6 was included into the framework. This device was designed to improve the grasping of thin and tiny objects by additionally considering pinches of the users' fingers for grasp detection. Therefore a DSO was developed that communicates with the pinch-sensitive hardware and reports the pinch states of all fingers to the grasp manager where this information is integrated into the grasping heuristics. The amount of user support by this hardware enhancement is evaluated in the study.

Tactile Feedback Two approaches for the integration of tactile feedback into the optical finger tracking devices have been presented within this work (see Section 4). These devices — one providing pressure-based feedback by using thin shape memory alloy wires and the other incorporating voice-coil motors as vibration actuators into the finger thimbles — were designed to support the user during the grasping process by providing grasp feedback. With this feedback the users should be able to immediately recognize when they have grasped an object. Up until this point, the amount of interaction support by tactile feedback remains debatable(see Section 2.3.2). It will be interesting to see if an interaction benefit can be proven by the user study.

Visual Grasping Feedback As a reference for the feedback devices, a visual grasping feedback was also integrated into the interaction framework. Therefore the grasping finger segments of the virtual hand representation were colored signal orange. This functionality was implemented as a DSO listening to the same grasp events used for the tactile feedback. This

Figure 5.4: Visual feedback in the CAVE, coloring virtual finger segments which are involved in a valid grasp.

visual feedback gives the users the chance to explicitly see whether they have grasped an object. For this kind of feedback no additional hardware, despite the original finger tracking system, is needed. Of course, due to the high precision of the optical glove used here, the virtual hand representation almost perfectly matches the real hand of the users. This becomes a problem for visual feedback when it is used in a projection-based system. Here, the real hand almost covers the virtual hand representation so that the visual feedback cannot be clearly seen. In the HMD setup, which completely immerses the user, visual feedback can be perceived easily by the users. However, since a simplified hand model was used with average finger and palm sizes, parts of the virtual hand can be seen in the CAVE-setup as well (Figure 5.4). As an alternative it would be possible to color the grasped objects instead of the hand geometry. However, this approach is not accepted by the users since it would significantly falsify the visual appearance of the virtual car.

Entire Prototype Application With these extensions the interaction framework described in Section 5.1.1 became more complex. It was extended to provide the proposed in- and output devices, the controller-based indirect interaction metaphor, the visual feedback and both display systems addressed within this work. An overview of the complete framework architecture can be seen in Figure 5.5.

Design

A common measure for the evaluation of interaction metaphors and their influence factors is to compare task completion times (TCT), as suggested by Zhai and Milgram [ZM98]. Direct interaction tasks can be separated into three phases: selection, manipulation and deselection.

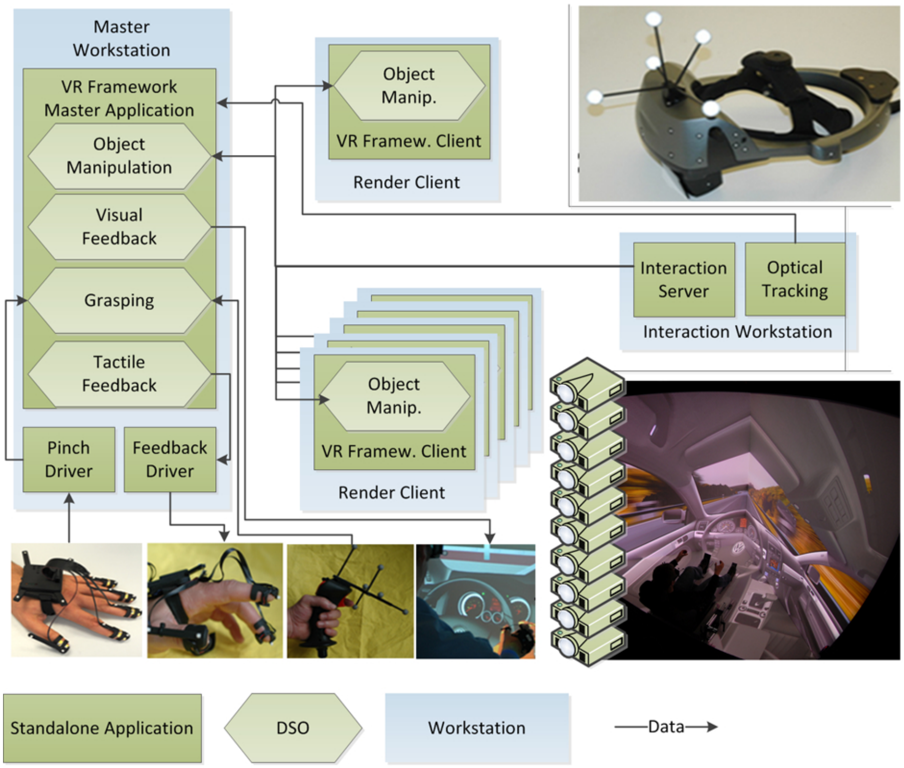

Figure 5.5: The interaction framework architecture.

In the context of direct finger-based interaction, these phases are realized by reliable grasping heuristics, a realistic and intuitive manipulation and reliable releasing of the object. The TCTs directly depend on the usability of the implementation of each task phase and thus represent a useful measure for the usability of the whole interaction metaphor.

However, finger-based interaction tasks are influenced by many factors that do not necessarily depend on the involved devices and metaphors, such as tracking reliability and individual user performance. Thus task performance cannot be the only measure to evaluate interaction. That is why, in this study, subjective judgments and the evaluation of user preferences help to complement the performance information provided by the TCTs.

Five tasks that had to be accomplished were defined for the study. Each of the tasks involved the following subtasks:

1. Grasp an object
2. Relocate the object to a pre-defined target position
3. Release the object

It was necessary that the study tasks were representative for common situations the users have to deal with in the desired applications. That is why a set of objects was chosen that represents

Table 5.1: Control Points and Tolerance of the Objects

Object	No. of Reference Points	Tolerance
Sun Shield	1	$20mm$
Driver's Door	1	$30mm$
Interior Mirror	4	$10mm$
Soda Bottle	2	$20mm$
Light Switch	1	$5mm$

the variety of common objects found in the car interior. For the user study the following five tasks were chosen:

1. The rotation of the driver's sun shield from design position to stop position
2. The opening of the driver's door until it reaches its stop position
3. The rotation of the interior mirror towards the user
4. The passing of a soda bottle from the passenger's footwell to the center console
5. The rotation of the light switch to its stop position

During the study, push interaction functionality was disabled so that each object had to be grasped explicitly. With the chosen objects, a broad range of object types typically found in cars is covered including several kinds of constraints and different object shapes, locations and sizes.

It was essential for the TCT-measurement to include the possibility to decide for task completion. Therefore, reference points for each manipulated object and target positions for these points were defined. The target position of an object is reached, if each of its reference points is within a tolerance region of the respective target reference points. Figure 5.6 shows the objects used within the evaluation scenario, their reference points and the size of the tolerance region. The reference points were visualized by small red spheres. The target position was shown with the help of semi-transparent colored duplicates of the objects and the tolerance regions were visualized by semi-transparent green spheres with an appropriate radius. The number of reference points per object and the respective size of the tolerance region can be seen in Table 5.1.

Visual feedback was integrated, informing the user about task completion. Therefore a green sphere appears around the virtual car, once the object reaches its target position. The start and stop condition for the measurements was the presence of the index finger tip within a semi-transparent purple sphere around the center of the steering wheel ($r = 80mm$). It became opaque when the condition was reached. Since there are objects that are manipulated with the left hand (door, light switch) or the right hand (sun shield, mirror, bottle) the corresponding index finger was chosen for the start-stop decision. Figure 5.7 shows a 2D screenshot of the complete user study scenario.

The subjects had to perform each of the five interaction tasks three times in sequence with each condition. The order of the objects was not varied since the differences among the objects

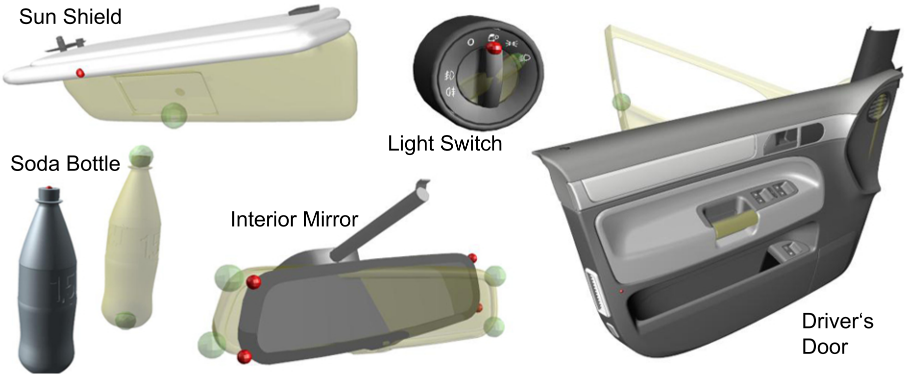

Figure 5.6: The objects used for evaluation, their semi-transparent target locations, the objects' reference points (red spheres) and the tolerance region (green spheres) used for task completion detection.

Table 5.2: Device and interaction properties questions

ID	Question
Q1	How unobtrusive was the hand-worn device?
Q2	Please rate the intensity of the feedback!
Q3	How comfortable was the feedback for you?
Q4	How well have you been able to grasp the object?
Q5	How well have you been able to recognize if you had grasped an object?
Q6	How well have you been able to release an object?
Q7	How immediately did you feel the feedback?
Q8	How unobtrusive was the feedback for you?
Q9	Please rate the realism of the interaction!
Q10	Have you been able to judge the car functionality?

have not been evaluated. The objects were reset to design position each time. The subjects were told to carefully place the objects at the target position. However, to fulfill the task, it was sufficient that the object reached the target position once for a short time. It was not necessary that it stayed in target position after releasing to avoid multiple grasp and release sequences.

After each interactional task — conducted three times in sequence — the subjects were asked to judge this particular interaction on a seven item Likert-scale with *one* being "very poor grasping, placing and releasing" and *seven* meaning "very good". It was interesting to see, if this subjective judgment would correlate with the task performance measures. After all interaction tasks of one condition the participants filled out a questionnaire with questions concerning individual device and interaction properties. The questions are listed in Table 5.2. Again, each of the questions had to be rated on the same seven item Likert-scale. Of course feedback-related questions ($Q2$, $Q3$, $Q7$ and $Q8$) only had to be answered if feedback was provided. After the subjects tried all conditions, the study was complemented with preference questions. Here the participants had to directly choose one of two options (Table 5.3).

If an interaction task was too difficult for the user in one condition, this particular task was

Figure 5.7: Screenshot of the interaction scenario while the user is interacting with the sun shield.

Table 5.3: User preference questions of the questionnaire

ID	Question
Q11	Do you prefer to use the pinch-sensitive device?
Q12	Which one do you prefer: the pressure-based or the vibration-based tactile feedback?
Q13	Which one do you prefer: visual or tactile feedback
Q14	Do you prefer to have any kind of feedback during interaction?
Q15	Do you prefer using the CAVE or the HMD?

canceled. In this case, the subjective rating is “one” and the highest (worst) TCT of this particular task of all other runs of all conditions was noted. Significance levels were calculated by a t-test with repeated measures. The global significance level of $p = 0.05$ was adjusted with the Holm-Bonferroni method [Hol79] for multiple tests.

Procedure

The experiment was conducted in the already described scenario that was also used for the expert review. There were twelve conditions that were evaluated, combining two display systems with six input devices (Table 5.4). Each of the twelve users performed the interaction tasks in each of the twelve conditions. To simplify the study process, the participants were divided into two groups. One group started evaluating the six input devices in the CAVE, while the other one started with the HMD. Each group continued the study with the other display system afterwards. The order of the devices was defined by a Latin square design (Table 5.5). In order to simplify the study flow and to reduce the number of calibrations and device changes the five-finger pinch-

Table 5.4: Conditions of the device-influence study

		CAVE (ANY.C)	HMD (ANY.H)
Finger Tracking Device	(FT)	FT.C	FT.H
Visual Feedback	(VF)	VF.C	VF.H
Pinch-sensitive Device	(PSD)	PSD.C	PSD.H
Pressure-based Tactile Feedback	(PTF)	PTF.C	PTF.H
Vibration-based Tactile Feedback	(VTF)	VTF.C	VTF.H
Flystick	(FL)	FL.C	FL.H

sensitive device was used for all FT-, PSD- and VF-conditions. For FT- and VF-conditions the pinch analysis was switched off and for VF-conditions visual feedback was added in the scene. Please note that the tactile feedback systems do not provide thimbles for the ring and pinky fingers, which play a minor role during grasping interaction. That is why the ring and pinky fingers were completely ignored for grasp decisions in any condition. Each of the devices was carefully calibrated to the individual user's hand and finger characteristics.

The subjects had to perform each of the five interaction tasks three times in sequence with each condition. Before they started with each trial, they were encouraged to familiarize themselves with the condition by interacting freely with all objects. After each interactional task the users were asked for their subjective judgment of this individual task. After all interaction tasks of one condition, the participants filled out the device properties questionnaire. After the subjects tried all conditions, the study was completed with preference questions. The complete course of actions can be seen in Table 5.6.

5.2.2 Results and Discussion

It was interesting to see whether or not the interaction would be influenced by the display type. The first subsection will discuss the study results with this focus. The second chapter compares the direct finger-based interaction with its benchmark method, the controller-based indirect technique. A pinch-sensitive mechanism incorporated into the finger tracking device is intended to increase the grasping capabilities of the users, the third subsection discusses if this expectation is fulfilled. The final subsection deals with grasping feedback. Here the users are informed by three different approaches that they succeeded in grasping an object and the subsection discusses if this leads to improved interaction parameters. For each aspect a number of theses are postulated which are verified or disproved by the results of the study.

The overall benchmark for the described tasks is the task performance in a real car interior. For the retrieval of informational reference values, a user in the very same real car was observed interacting with the very same objects. Of course the target positions differed slightly and the times had to be captured manually. The user had to touch the badge at the steering wheel for initiating and finishing the interaction. Each interaction was performed three times and TCTs were averaged (Table 5.7).

Table 5.5: Sequence of the conditions for each participant

Subject	1	2	3	4	5	6	7	8	9	10	11	12
$No.1$	HMD						CAVE					
	FT	VF	PSD	PTF	VTF	FL	FT	VF	PSD	PTF	VTF	FL
$No.2$	HMD						CAVE					
	FL	FT	VF	PSD	PTF	VTF	FL	FT	VF	PSD	PTF	VTF
$No.3$	HMD						CAVE					
	VTF	FL	FT	VF	PSD	PTF	VTF	FL	FT	VF	PSD	PTF
$No.4$	HMD						CAVE					
	PTF	VTF	FL	FT	VF	PSD	PTF	VTF	FL	FT	VF	PSD
$No.5$	HMD						CAVE					
	PSD	PTF	VTF	FL	FT	VF	PSD	PTF	VTF	FL	FT	VF
$No.6$	HMD						CAVE					
	VF	PSD	PTF	VTF	FL	FT	VF	PSD	PTF	VTF	FL	FT
$No.7$	CAVE						HMD					
	FT	VF	PSD	PTF	VTF	FL	FT	VF	PSD	PTF	VTF	FL
$No.8$	CAVE						HMD					
	FL	FT	VF	PSD	PTF	VTF	FL	FT	VF	PSD	PTF	VTF
$No.9$	CAVE						HMD					
	VTF	FL	FT	VF	PSD	PTF	VTF	FL	FT	VF	PSD	PTF
$No.10$	CAVE						HMD					
	PTF	VTF	FL	FT	VF	PSD	PTF	VTF	FL	FT	VF	PSD
$No.11$	CAVE						HMD					
	PSD	PTF	VTF	FL	FT	VF	PSD	PTF	VTF	FL	FT	VF
$No.12$	CAVE						HMD					
	VF	PSD	PTF	VTF	FL	FT	VF	PSD	PTF	VTF	FL	FT

Table 5.6: Course of actions of the user study with trials and measures

ID	action	measure
0	Statistical information	questionnaire
1	Display1	
1.0	Familiarization with the display	
1.1	Condition1	
1.1.0	Familiarization with the condition	
1.1.1.1	Task1, Trial1	TCT
1.1.1.2	Task1, Trial2	TCT
1.1.1.3	Task1, Trial3	TCT, Subjective
1.1.2.1	Task2, Trial1	TCT
...		
1.1.2.3	Task2, Trial3	TCT, Subjective
...		
1.1.5.3	Task5, Trial3	TCT, Subjective
1.1.6	Judgment of condition1	Device questionnaire
1.2	Condition2	
...		
1.6.6	Judgment of condition6	Device questionnaire
2	Display2	
...		
2.6.6	Judgment of condition6	Device questionnaire
3	Retrospective technology judgment	Preference questions

Table 5.7: Reference interaction times for the five objects in a real car interior

Object	Mean TCT (ms)	SD
Sun Shield	1627	106.93
Driver's Door	2870	166.43
Interior Mirror	2247	144.68
Soda Bottle	2827	137.96
Light Switch	1377	5.77

Influences of Display Device

The first interesting aspect that was analyzed is the influence of display type on user interaction. From the everyday experience in industrial applications differences resulting from the individual characteristics of the devices have to be expected.

The three-sided CAVE has a very high field of view almost covering the whole human field of view. The head-worn hardware necessary for image separation and the stereoscopic effect is lightweight and unobtrusive. The presence of the users' own body provides a size reference that could increase the users' confidence with the size of the perceived virtual objects. The head mounted display is fully immersive since it completely hides reality from the users, the only self reference is the virtual hand. Consequently — in contrast to projection-based systems — the localization of the users' hands at the virtual objects is not influenced by focus shifts. On the other hand, the everyday experience shows that users of HMD-applications often complain on incorrect proportions.

Considering these aspects two effects can be expected that should have an influence on task performance and subjective user judgment:

Thesis 1 *Interaction in the CAVE is preferred by the users, leading to higher subjective rates for interaction in this environment.*

Thesis 2 *Interaction in the HMD is more efficient and reliable because of the fully immersive character.*

User Preference and User Observation Display system preference was one of the user preference questions of the questionnaire (Table 5.3). Surprisingly, user preference was quite balanced. After interacting with the direct finger-based interaction scenario, seven users preferred the Head Mounted Display, while five preferred the CAVE. Since the users had the chance to directly compare both systems, they were able to perceive the individual advantages and disadvantages of both types of display. HMDs fully exclude the real world and thus avoid the typical focus and occlusion problems resulting from the interaction between real hands and virtual objects. It can be assumed that the HMD application strongly benefits from direct finger-based interaction, because users have a self-reference — their virtual hands — in the virtual world, which has been shown to improve size perception [RIKA08]. Thus it is no surprise that

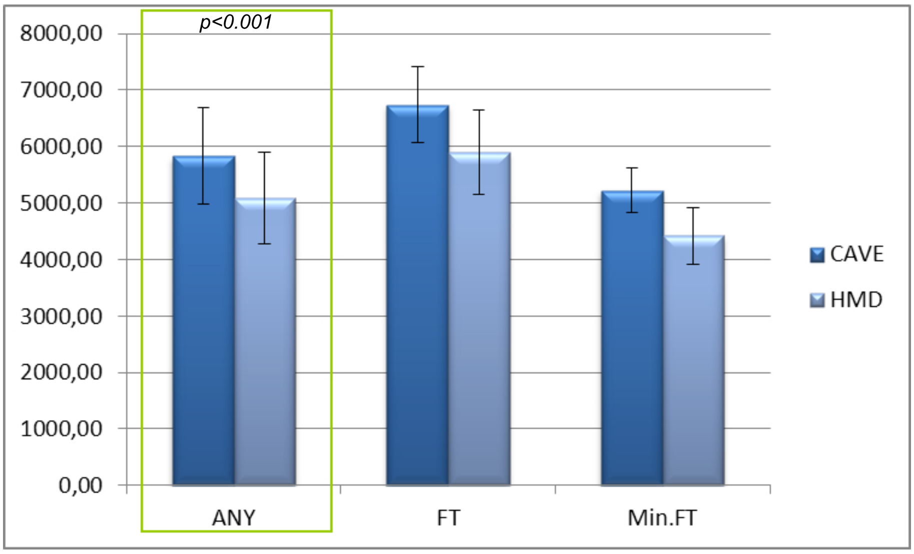

Figure 5.8: Task completion times in ms in the CAVE vs. the HMD in all conditions (ANY), the bare finger tracking condition (FT) and the fastest fingertracking TCTs (Min.FT). Green frames indicate significant effects.

users did not complain about incorrect size perception of the car interior. However some of the users criticized that the narrow field of view (80° diagonal) of the HMD made it necessary to extensively move the head during interaction. Although this particular display is relatively comfortable the increased discomfort and fatigue were mentioned by some of the users. Interestingly some users found it difficult to judge the distance of virtual objects in the HMD condition. No user complained about the virtual hands penetrating virtual objects during grasps.

Task Performance Besides the direct user preference it was interesting to see if one of the display devices would allow for a significant advantage for the users of a direct finger-based interaction scenario. Therefore mainly the bare finger tracking conditions are of interest, because they are free of further influences. The task completion times of the bare finger tracking conditions alone did not show a significant advantage for any of the displays (*FT.C* vs. *FT.H*). However, a tendency towards faster interaction in the HMD condition can be observed if the fastest TCTs for each task are compared (*MIN(FT.C)* vs. *MIN(FT.H)*). Moreover, comparing the mean task completion times of all conditions — including all hand-hardware conditions — in the HMD with those in the CAVE a significant advantage for the HMD can be seen (*ANY.C* vs. *ANY.H*: $t = 3.9$, $p = 0.0002$, $df = 71$) (Figure 5.8).

Subjective Measures The results of the subjective interaction judgments were comparable in CAVE and HMD condition for the bare finger tracking conditions as well as comparing all conditions (*FT.C* vs. *FT.H* and *ANY.C* vs. *ANY.H*). Having a closer look at the bare finger tracking conditions, one object was judged considerably better for each display system — the

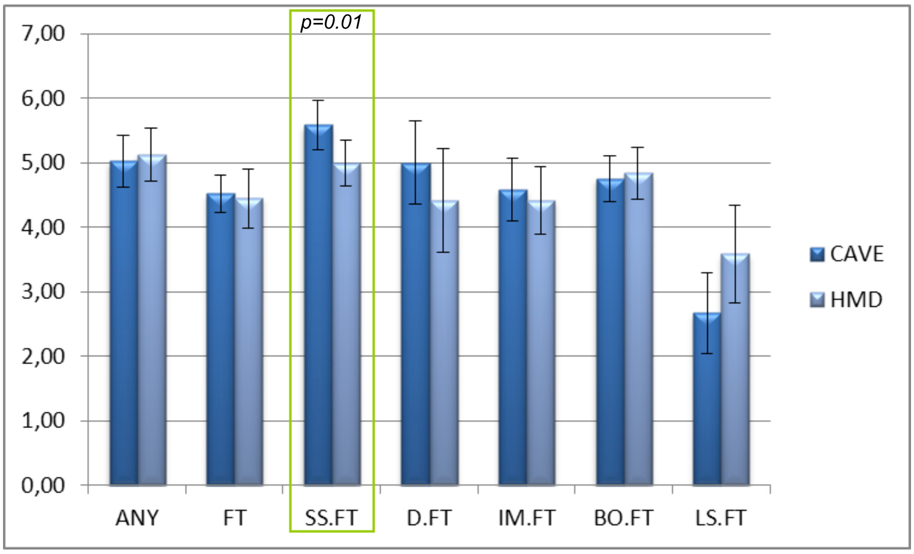

Figure 5.9: Subjective measures (0=very bad,...,7=very good) in the CAVE vs. the HMD for all conditions (ANY), the bare finger tracking (FT) and the finger tracking conditions split into the respective objects, sun shield (SS.FT), door (D.FT), interior mirror (IM.FT), bottle (BO.FT), and light switch (LS.FT). Green frames indicate significant effects.

sun shield interaction with a significant effect under CAVE-condition (*SS.FT.C* vs. *SS.FT.H*: $t = -3.0$, $p = 0.01$, $df = 11$) and the light switch without significance, when an adjusted significance-level is used, under HMD-condition (*LS.FT.C* vs. *LS.FT.H*: $t = 2.4$, $p = 0.03$, $df = 11$) (Figure 5.9). This complies with the expectation of display influence since the sun shield interaction does not require a precise judgment of the hand position in relation to the sun shield. In contrast, the spatial relation of the user's hand to the tiny light switch is more easily perceivable in the purely virtual environment of the HMD.

Concerning the short questionnaire, again both systems were rated comparably in the bare finger tracking conditions (Figure 5.10). However, a tendency for better values for CAVE-interaction was apparent for the bare finger tracking condition (Figure 5.11). It was interesting to see that for both systems, interaction realism and the ability to judge functional aspects of a car were said to be above average (*Q9.FT.C*: 5.08, *Q9.FT.H*: 4.58, *Q10.FT.C*: 4.83, *Q10.FT.H*: 5.08). Surprising was that for the subjects the HMD-application was more appropriate for functional car assessments than the CAVE-environment. The advantage of the HMD is not a significant effect (*Q10.FT.C* vs. *Q10.FT.H*: $t = -0.56$, $p = 0.59$, $df = 11$) but if anything, a significant effect in favor of the CAVE could have been expected, due to the reduced FOV and comfort of the HMD.

Concerning question $Q10$ another interesting display-related effect could be seen. The difference between direct finger-based interaction and controller-based indirect interaction in the CAVE was not as considerable as it was in the HMD and thus was not significant (*Q10.FL.H* vs.

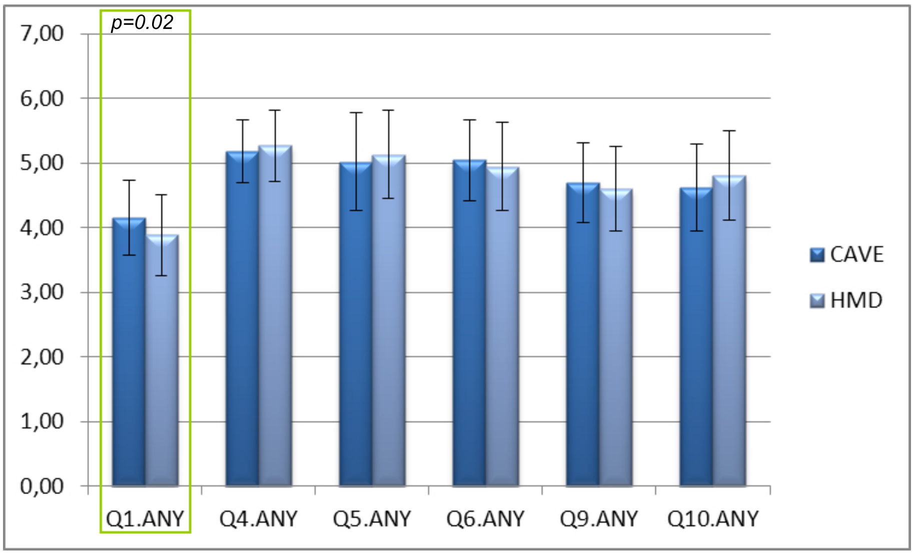

Figure 5.10: Subjective measures of the questionnaire (0=very bad,...,7=very good) in the CAVE vs. the HMD for all conditions (ANY), concerning the hand device obtrusiveness (Q1), grasping capabilities (Q4), grasping judgment (Q5), releasing capabilities (Q6), interaction realism (Q9) and car judgment (Q10). Green frames indicate significant effects.

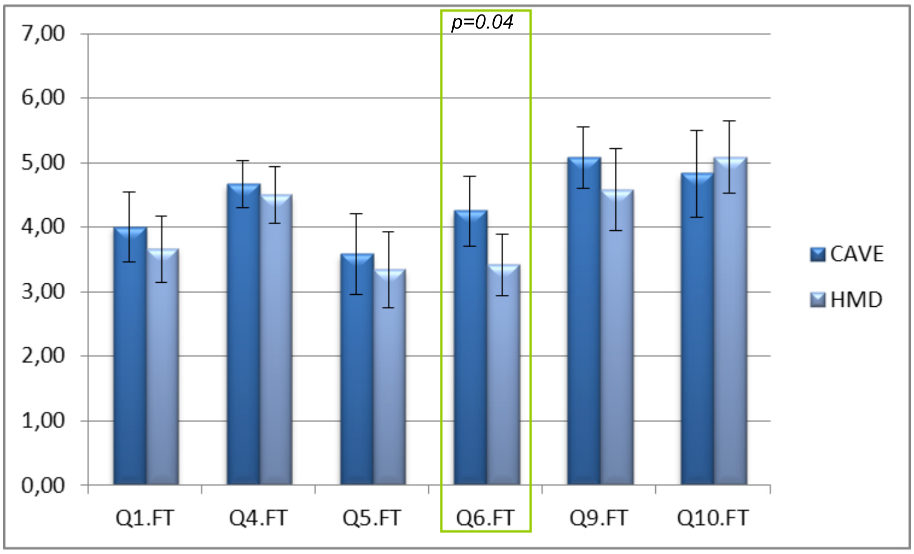

Figure 5.11: Subjective measures of the questionnaire (0=very bad,...,7=very good) in the CAVE vs. the HMD for the bare finger tracking condition (FT), concerning the hand device obtrusiveness (Q1), grasping capabilities (Q4), grasping judgment (Q5), releasing capabilities (Q6), interaction realism (Q9) and car judgment (Q10). Green frames indicate significant effects.

Q10.FT.H: $t = -3.2$, $p = 0.009$, $df = 11$ and *Q10.FL.C* vs. *Q10.FT.C*: $t = -1.03$, $p = 0.32$, $df = 11$). This display-related effect was to be expected. The HMD completely excludes the real world, that is why the users only see the virtual cursor. Using a projection-based system, the input device in their hand and the virtual representation as an extension into the virtual environment are visible. Consequently, the users are able to judge functional aspects of the virtual car such as accessibilities even with this abstract metaphor. However, it is clear that more reliable decisions can be made with the realistic direct interaction technique.

Another noticeable display-related effect occurred with the visual feedback conditions. This kind of feedback succeeds differently in both displays. This results from the special character of the feedback implementation and will be discussed in a separate section.

Conclusion Surprisingly, the first thesis for this section has to be rejected, as the CAVE-like display system was not preferred by the users (Thesis 1). The second thesis, postulating that interaction in the HMD is more efficient and reliable, was only partly confirmed by a slightly better task performance (Thesis 2). In general the tests indicate that for functional assessments of the car interior direct finger-based interaction can be equally well used in both display systems — the CAVE and the HMD. The decision regarding which system should be used can be made based on application-dependent factors, such as the requirements on passive haptics or the size of the manipulated objects.

Direct vs. Indirect Interaction

One of the central questions resulting from the realization of direct finger-based interaction is, how this interaction technique performs compared to the common quasi-standard — indirect controller-based interaction. This ray-based picking metaphor utilizing an input device has become very popular especially in industrial immersive VR-applications due to its simplicity and ease-of-use. The explicit interaction trigger by pressing and releasing a button on the controller leads to highly robust and reliable “grasping” and an interaction status that is always transparent to the user. Unfortunately, the means to reach this robustness is abstraction and the indirect metaphor is not comparable to the real human-car interaction. Consequently, functional aspects of the car, such as the accessibility and usability of interior elements, hardly can be analyzed with this interaction metaphor. This was the motivation for the implementation of direct finger-based interaction. This section shows, if and to what extent the indirect controller-based method is more efficient and if — in contrast — the direct finger-based technique enables the users to judge car functionality. The experience with both interaction metaphors leads to three central theses:

Thesis 3 *Indirect controller-based interaction is more robust and efficient than direct finger-based interaction, leading to shorter task completion times and increased subjective user judgments.*

Thesis 4 *This increase of robustness and efficiency is mainly caused by robust and reliable “grasping” leading to higher subjective grasping rates.*

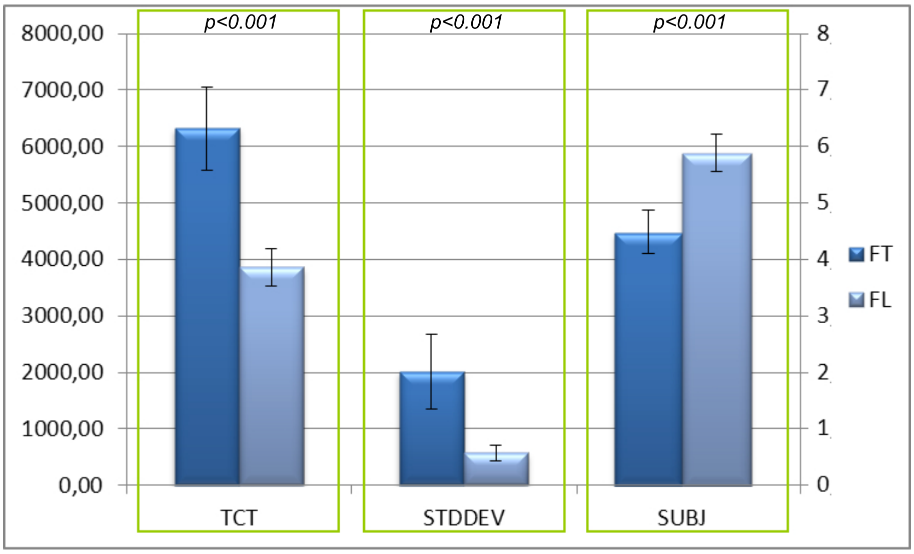

Figure 5.12: Task performance in ms (TCT), intra-subjective standard deviations in ms (SD) and subjective interaction judgment (SUBJ) (0=very bad,...,7=very good) for the bare finger tracking (FT) and the Flystick condition (FL). Please note, that the subjective measure refers to the right-hand ordinate. Green frames indicate significant effects.

Thesis 5 *Direct finger-based interaction is more realistic than the indirect method and consequently enables the users to judge the functionality of virtual cars in immersive applications.*

Task Performance and Subjective Judgments The study clearly confirmed the expectations. The indirect Flystick-based metaphor performed considerably better than unsupported direct interaction using the bare finger tracking. Task completion times significantly differed and were considerably faster for all Flystick interactions (*FL* vs. *FT*: $t = -8.64$, $p < 0.001$, $df = 23$). This quantitative benefit is also reflected by subjective interaction ratings (*FL* vs. *FT*: $t = 8.91$, $p < 0.001$, $df = 23$). Furthermore, the interaction with the Flystick was more robust leading to significantly lower intra-subject standard deviations of task performance (*FL* vs. *FT*: $t = -4.98$, $p < 0.001$, $df = 23$) (Figure 5.12).

Questionnaire As was expected, the advantage of the indirect interaction is that it makes “grasping” very easy — only pointing and pressing a button is required. It is evident that the very fast and subjectively better Flystick interaction correlated with a significantly higher subjective rating for all grasping questions (*Q4.FL* vs. *Q4.FT*: $t = 10.122$, $p < 0.001$, $df = 23$, *Q5.FL* vs. *Q5.FT*: $t = 8.26$, $p < 0.001$, $df = 23$ and *Q6.FL* vs. *Q6.FT*: $t = 12.87$, $p < 0.001$, $df = 23$).

Although higher task performance and user ratings can be achieved with indirect Flystick-based interaction, it has to be expected that a realistic interaction needs direct input. This also was confirmed by the questionnaire. The participants rated controller-based interaction consid-

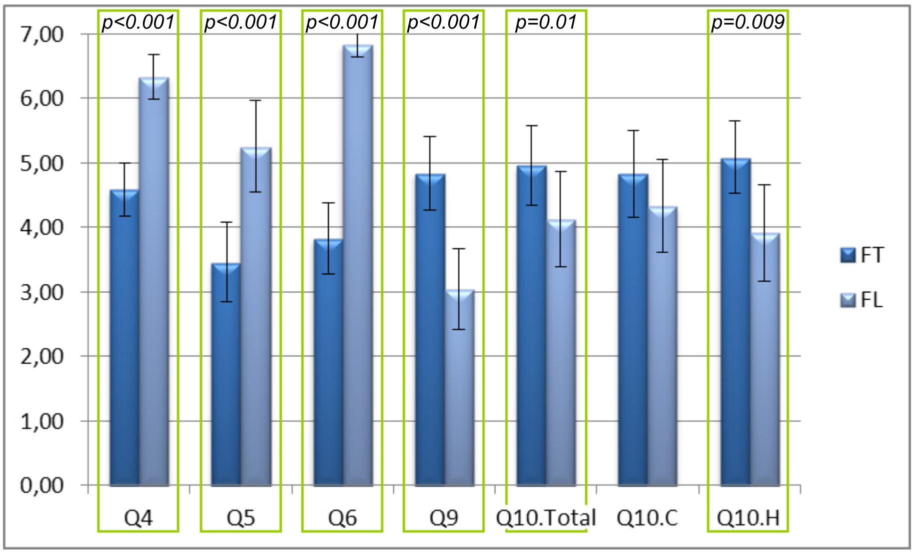

Figure 5.13: Subjective ratings of the questionnaire (0=very bad,...,7=very good) for the bare finger tracking (FT) and the indirect FLystick condition (FL), concerning the grasping capabilities (Q4), grasping judgment (Q5), releasing (Q6), interaction realism (Q9) and car judgment (Q10). For question 10 the whole condition (Q10.Total) is split related to the displays (Q10.C and Q10.H). Green frames indicate significant effects.

erably less realistic (*Q9.FL* vs. *Q9.FT*: $t = -5.84$, $p < 0.001$, $df = 23$). While the Flystick interaction has to be considered to be an abstraction, direct interaction provides a comparatively more realistic interaction. With higher levels of realism, the properties of a virtual car interior can be evaluated better. Consequently, subjective ratings for the question on the ability to judge the virtual car are higher for direct interaction (*Q10.FL* vs. *Q10.FT*: $t = -2.7327$, $p = 0.012$, $df = 23$). As previously discussed, a display-related difference was noticeable. The difference in Q10-ratings were not significant in the CAVE conditions alone (*Q10.FL.C* vs. *Q10.FT.C*) (Figure 5.13).

Conclusion All expectations were fully confirmed by the study and consequently all three theses (Theses 3, 4 and 5) postulated beforehand were fully proven. The indirect controller-based interaction method provides more robust and reliable selection and thus it is more efficient than direct finger-based interaction. On the other hand, the finger-based method provides higher realism and enables the users to judge a car virtually.

Increased Grasping Capabilities with the Pinch-sensitive Device

The pinch-sensitive finger tracking device is able to recognize pinches between the real fingers of the users. Pinches occur quite often during interaction, especially when tiny or thin objects are grasped. With the pinch-sensitive devices this information is used to improve grasp detection. Therefore, the grasping heuristics is extended as it was discussed in Chapter 3.1.6. The users were not informed about the special character of this input device before the trials. In this section, the pinch-sensitive device is compared with the bare finger tracking.

The pinch-sensitive finger tracking device was designed to improve the grasping capabilities of the users. At the same time it still provides the realistic direct finger-based interaction as it was proposed within this work. Consequently the following effects have to be expected:

Thesis 6 *The pinch-sensitive system increases the users' grasping capabilities which is represented by higher ratings compared to the bare finger tracking for the grasping-related questions of the questionnaire.*

Thesis 7 *Better grasping capabilities also lead to higher task performance, subjective ratings and interaction robustness, compared to bare finger tracking.*

Thesis 8 *Since the realistic character of direct finger-based interaction is maintained with this device, realism and car judgment capabilities are rated comparable to the standard technique.*

Thesis 9 *Finger pinches mainly occur when thin and tiny objects are grasped, that is why an advantage can be seen especially for the challenging objects such as the door handle and the light switch.*

Task Performance and Subjective Judgments As expected, a clear interactional improvement can be obtained with the pinch-sensitive device (Figure 5.14). The achieved task completion times are significantly better than those of the bare finger tracking (*PSD* vs. *FT*: $t = -6.11$, $p < 0.001$, $df = 23$). They are only 10 to 20% — yet significantly — slower than those of the indirect Flystick-based interaction (*PSD* vs. *FL*: $t = 4.02$, $p < 0.001$, $df = 23$) and come closer to the task performance encountered in reality (Figure 5.15).

Again the higher task performance is reflected by the subjective ratings given by the subjects after each interaction (*PSD* vs. *FT*: $t = 4.97$, $p < 0.001$, $df = 23$). As with indirect interaction a higher robustness was indicated by lower intra-subject standard deviations of task performance compared to the bare finger tracking (*PSD* vs. *FT*: $t = -4.37$, $p < 0.001$, $df = 23$).

Interestingly the benefit for task performance of the pinch-sensitive system could not be equally found for all interaction tasks (Figure 5.16). Two of the objects are quite voluminous (interior mirror, soda bottle). Usually they are not grasped with a pinch grasp and consequently the pinch detection was not activated when grasping them. Thus, pinch detection did not improve grasp performance (*mirror.PSD* vs. *mirror.FT*: $t = -1.09$, $p = 0.29$, $df = 23$ and *bottle.PSD* vs. *bottle.FT*: $t = -0.92$, $p = 0.37$, $df = 23$). Grasping for the other three objects was facilitated by pinch detection since the subjects almost always employed a pinch grasp (*door.PSD* vs. *door.FT*: $t = -4.56$, $p < 0.001$, $df = 23$ and *light.PSD* vs. *light.FT*: $t = -5.87$, $p < 0.001$, $df = 23$ and *shield.PSD* vs. *shield.FT*: $t = -2.16$, $p = 0.04$, $df = 23$). Especially the light

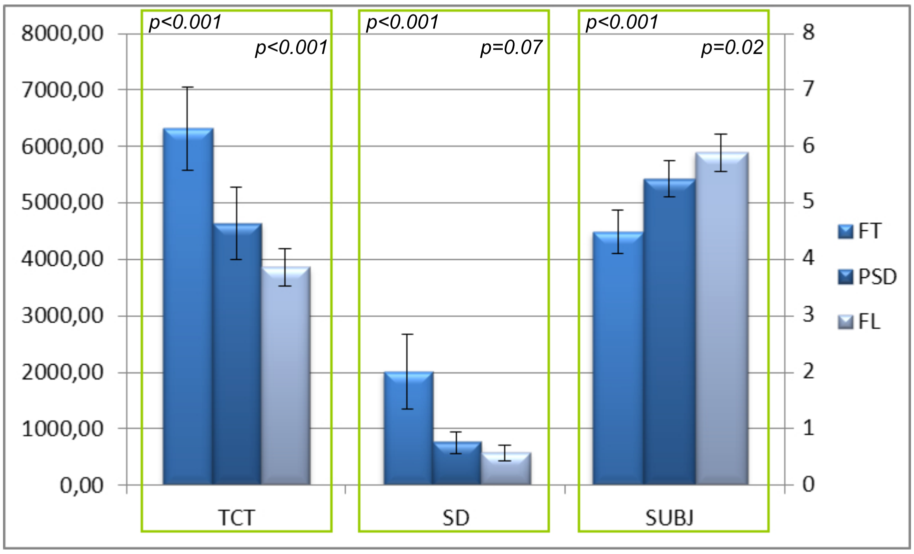

Figure 5.14: Task performance in ms (TCT), intra-subjective standard deviations in ms (SD) and subjective interaction judgment (SUBJ) (0=very bad,...,7=very good) for the bare finger tracking (FT), the pinch-sensitive device (PSD) and the Flystick condition (FL). Please note, that the subjective measure refers to the right-hand ordinate. Green frames indicate significant effects. The left significance value within the frames refers to the difference between FT and PSD, the right value refers the difference of FT and FL condition.

switch which was, due to its minor size a problem for most of the users, could be grasped considerably better with pinch detection. Please note, that the difference for the sun shield does not prove significance if the significance level is adjusted with the Holm-Bonferroni method. This reflects the fact, that the sun shield could be easily grasped by the users in all conditions.

Questionnaire Obviously the pinch-sensitive device helps the users during direct interaction by significantly improving grasping. This also was approved by the results of the questionnaire (Figure 5.17). The users were able to grasp objects considerably better (*Q4.PSD* vs. *Q4.FT*: $t = 6.58$, $p < 0.001$, $df = 23$). It was surprising to see that the users even felt significantly supported with grasping judgment and object release (*Q5.PSD* vs. *Q5.FT*: $t = 4.98$, $p < 0.001$, $df = 23$ and *Q6.PSD* vs. *Q6.FT*: $t = 3.09$, $p = 0.005$, $df = 23$). Both issues are not explicitly supported by the device, especially releasing follows the same rules as they are used for the bare finger tracking. It might be the case that the overall better interaction with the pinch-sensitive device leads to this inexplicable improvement.

For the focussed applications, task performance and user judgment improvements can only be a benefit if the realism of the interaction metaphor and the ability to virtually evaluate a car interior are rated on a high level, at least comparable to the bare finger tracking device. In fact, the PSD-conditions are rated considerably higher than the Flystick-conditions for Q9 and Q10

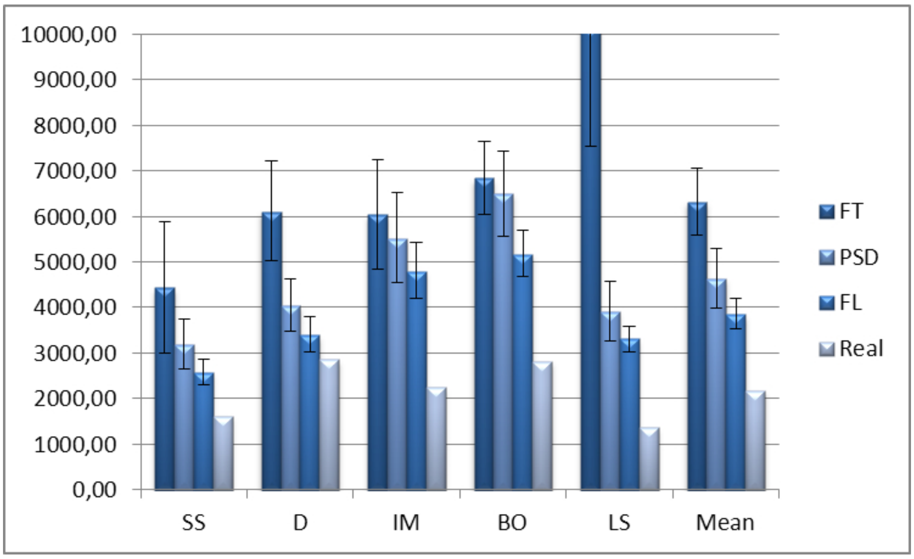

Figure 5.15: Task performance in ms for the bare fingertracking (FT), the pinch-sensitive device (PSD), the Flystick-condition (FL) and the trial in the real car (Real), with respect to the interaction objects, sun shield (SS), driver's door (D), interior mirror (IM), bottle (BO) and light switch (LS). Error bars represent standard deviation. Please note, that the real trial was conducted with one subject only.

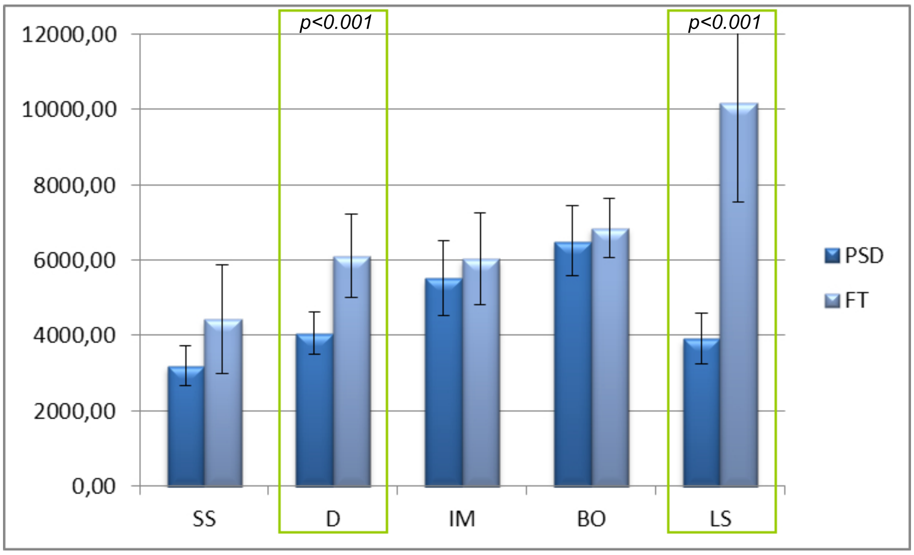

Figure 5.16: Task performance in ms for the pinch-sensitive device (PSD) vs. the bare finger tracking (FT) with respect to the interaction objects, sun shield (SS), driver's door (D), interior mirror (IM), bottle (BO) and light switch (LS). Green frames indicate significant effects.

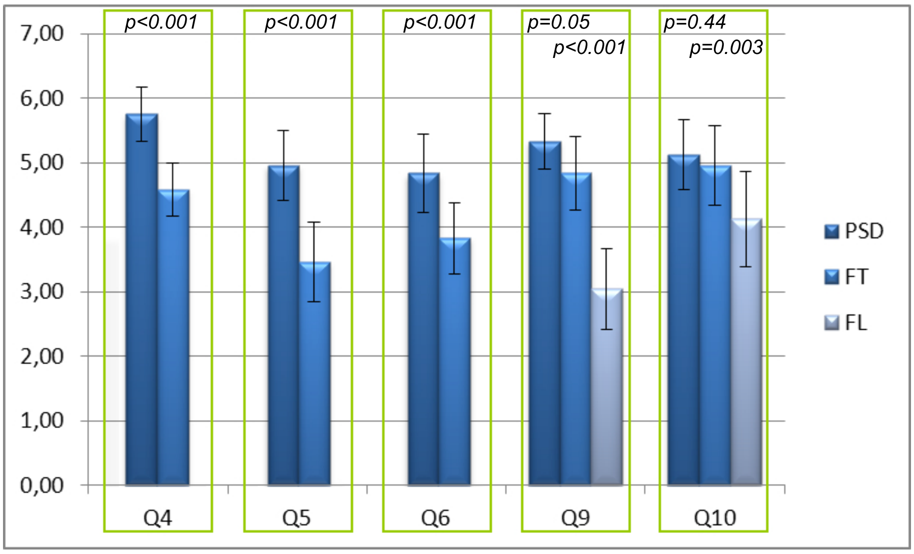

Figure 5.17: Subjective ratings of the questionnaire (0=very bad,...,7=very good) for the pinch-sensitive device (PSD), the bare finger tracking (FT) and the indirect FLystick condition (FL), concerning the grasping capabilities (Q4), grasping judgment (Q5), releasing (Q6), interaction realism (Q9) and car judgment (Q10). Green frames indicate significant effects. The left significance value within the frames refers to the difference between PSD and FT, the right value refers the difference of PSD and FL condition.

(*Q9.PSD* vs. *Q9.FL*: $t = 8.03$, $p < 0.001$, $df = 23$ and *Q10.PSD* vs. *Q10.FL*: $t = 3.32$, $p = 0.003$, $df = 23$). The user ratings on realism were even significantly better with the pinch-sensitive device compared to the bare finger tracking (*Q9.PSD* vs. *Q9.FT*: $t = 2.08$, $p = 0.049$, $df = 23$). The overall better Q9- and Q10-performance of the pinch-sensitive device compared to the interaction with bare finger tracking clearly shows that robustness and the ease of use are important for realistic interaction metaphors.

Conclusion Concerning the pinch-sensitive device all expectations were completely fulfilled and the four theses postulated beforehand were proved. The pinch-sensitive finger tracking device improves the grasping capabilities of the users — especially for those objects that are difficult to grasp (Theses 6 and 9). This improved grasping directly leads to higher task performance and a better subjective experience (Thesis 7). In contrast to the controller-based interaction metaphor this interaction improvement does not come along with reduced realism, because no abstraction of the actual interaction metaphor is introduced (Thesis8). This supporting technique enables functional validation of a virtual car together with robust and reliable grasping.

Improved Grasping Judgment by Grasping Feedback

Independent to the individual grasping capabilities, direct finger-based interaction can be supported by providing grasping feedback. Without a feedback, only the beginning motion of a grasped object is an indication for grasp evolution. The users do not recognize their grasping success immediately in this case. In contrary, the real world lets the users feel if they have grasped an object via the haptic channel.

Until now, it is not possible to provide appropriate force feedback in a way that the users can freely interact with virtual objects and feel these objects haptically. However, this makes this sense free to be stimulated by feedback methods that provide information on the grasping process. This can be reached by specialized hand hardware as was described in Chapter 4. This hardware presents grasping information to the users by stimulating pressure or vibration sensors within the skin of the users' finger tips. Especially the pressure based stimulation is quite close to the sensation users feel when they really touch an object.

As a reference for the tactile methods a visual grasping feedback also was implemented. Therefore the color of the virtual hands' finger segments that define a valid grasp is changed to signal orange. The disadvantage of this approach is that the virtual hand mostly disappears behind the real hand in projection-based virtual environments. However, it can be easily integrated into any application without integrating specialized hardware, as it is necessary for the tactile feedback. Furthermore, it does not affect the visual appearance of the virtual car as it is done by feedback approaches changing object properties. The unpersuaded impression of the virtual car is a key requirement of automotive users especially for design-related applications.

This part of the study was intended to reveal whether the grasping feedback would increase the grasping capabilities of the users and if this would lead to increasing task performances and improved subjective ratings. Furthermore it was interesting how the users perceived the individual characteristics of the different feedback approaches.

Visual Grasping Feedback The finger segments involved in a valid grasp were colored orange for the visual feedback. It had to be expected that this simple and comfortable feedback approach already is sufficient to achieve an interaction improvement. Due to the occlusion of the virtual hand in the projection-based environment, significant differences between the CAVE- and HMD-conditions have to be expected. To provide visual feedback by coloring the virtual objects is no alternative because it would destroy the visual appearance of the car. Without the necessity of additional hardware higher comfort ratings of this feedback type compared to tactile feedback should be achieved. Consequently, the following theses are postulated beforehand:

Thesis 10 *The visual feedback supports the users with grasping and therefore leads to better task completion times and subjective measures.*

Thesis 11 *The feedback works better in the HMD-condition since it is not masked by the users' real hands.*

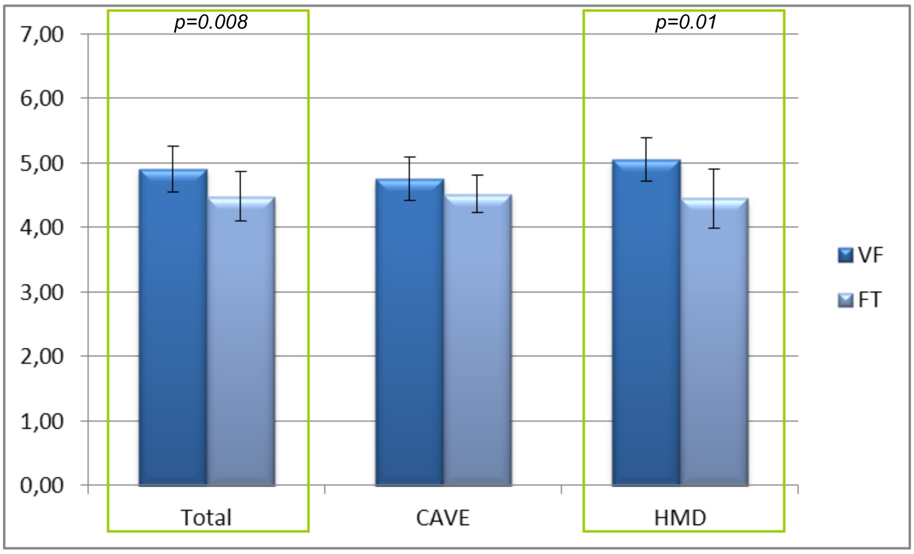

Figure 5.18: Subjective interaction judgment (0=very bad,...,7=very good) in the CAVE vs. the HMD for the visual feedback condition (VF) and the bare finger tracking (FT), disregarded the display (Total) and for the two display conditions (CAVE, HMD). Green frames indicate significant effects.

Thesis 12 *The visual feedback is more comfortable than the other feedback conditions since it does not rely on additional hardware.*

The study results reveal that visual feedback is only supporting the user subjectively (*VF* vs. *FT*: $t = 2.9$, $p = 0.008$, $df = 23$). And, as it had to be expected, this improvement can only be proven for the HMD-condition (*VF.C* vs. *FT.C*: $t = 1.13$, $p = 0.28$, $df = 23$ and *VF.H* vs. *FT.H*: $t = 3.12$, $p = 0.0097$, $df = 23$) (Figure 5.18).

Further explanations for the differences related to the display are provided by the results of the questionnaire (Figure 5.19). The visual feedback is perceived stronger in the HMD as it is in the CAVE (*Q2.VF.C* vs. *Q2.VF.H*: $t = 2.28$, $p = 0.043$, $df = 11$) and the support with grasp judgment is significantly higher (*Q5.VF.C* vs. *Q5.VF.H*: $t = 3$, $p = 0.012$, $df = 11$). Compared to the bare finger tracking, the visual feedback provides this intended improvement for the whole sample (*Q5.VF* vs. *Q5.FT*: $t = 4.6$, $p < 0.001$, $df = 23$) and in the HMD-condition (*Q5.VF.H* vs. *Q5.FT.H*: $t = 5.53$, $p < 0.001$, $df = 11$); for the CAVE-condition no significant effect was separately encountered (*Q5.VF.C* vs. *Q5.FT.C*: $t = 1.78$, $p = 0.10$, $df = 11$). A weak support of the feedback can be also found for the actual grasping (*Q4.VF* vs. *Q4.FT*: $t = 2.23$, $p = 0.036$, $df = 23$). This effect did not remain significant when the significance level was adjusted with the Holm-Bonferroni method for multiple tests. And it was not proven in the projection-based environment.

Since no additional hardware besides the finger tracking device is necessary, the feedback comes along with relatively high comfort ratings. The subjects rated the unobtrusiveness of the

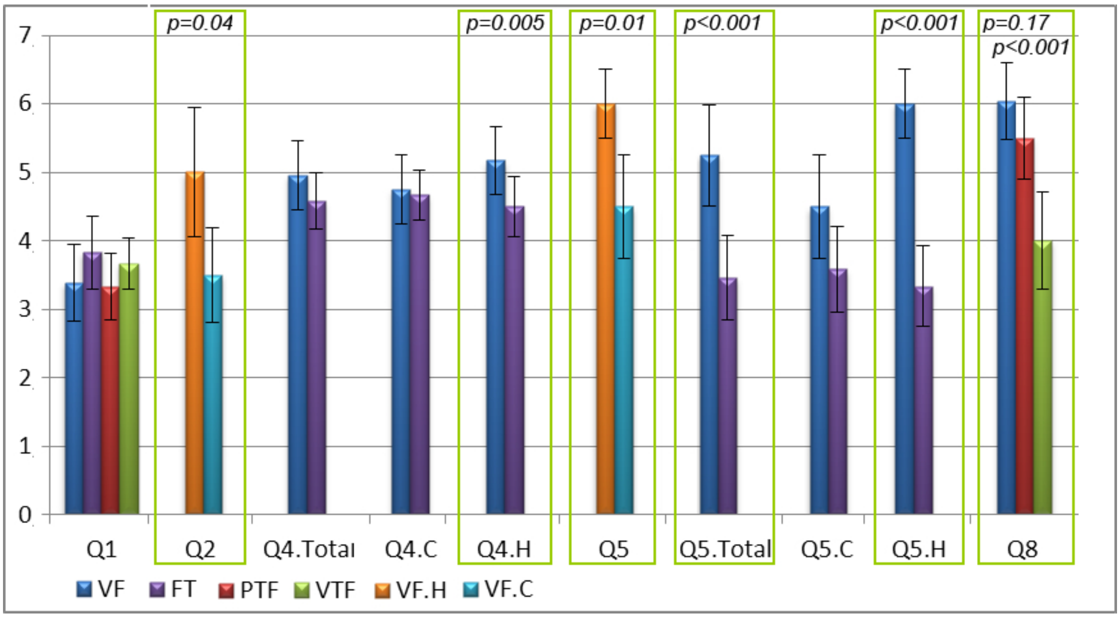

Figure 5.19: Subjective ratings of the questionnaire (0=very bad,...,7=very good) for the visual feedback condition (VF), the bare finger tracking (FT), the pressure-based (PTF) and the vibrational feedback device (VTF), concerning the hand device obtrusiveness (Q1), the strength of the feedback (Q2), the grasping capabilities (Q4), grasping judgment (Q5) and the obtrusiveness of the feedback (Q8). For Q2 and Q5 the visual feedback condition is compared with respect to the display type (VF.C and VF.H). For Q5 the VF-condition is compared with FT disregarded the display type (Q5.Total) and with respect to the used display type (Q5.C and Q5.H). Green frames indicate significant effects.

hardware higher than for the PTF conditions (*Q1.VF* vs. *Q1.PTF*: $t = 2.12$, $p = 0.045$, $df = 23$). The higher ratings compared to the vibration-based feedback did not yield a significant effect. In the next section, the difference of PTF and VTF concerning the comfort ratings are further evaluated. Compared to the bare finger tracking, no difference can be observed, which is reasonable, considering that the very same device is used. The unobtrusiveness of the feedback was rated higher for VF-conditions compared to the tactile feedback as well. Here, only the comparison to the vibration-based feedback had a significant effect (*Q8.VF* vs. *Q8.PTF*: $t = 1.41$, $p = 0.17$, $df = 23$ and *Q8.VF* vs. *Q8.VTF*: $t = 4.68$, $p < 0.001$, $df = 23$).

As it had to be expected, these results verify that a visual feedback at the virtual hand is only sufficient if a Head Mounted Display is used (Thesis 11). Nevertheless, it supports the users with grasping, leading to higher subjective judgments, as it was expected beforehand (Thesis 10). In contrast, this support did not lead to higher task performance for any of the conditions. The benefit of the feedback for actual grasping support (question $Q4$) is lower than for grasping judgment (question $Q5$). This is due to the fact that grasping feedback was provided, that informs the user when a valid grasp evolved, instead of giving collision feedback that eventually would help the user to actually grasp an object. The visual feedback was the most comfortable feedback method (Thesis 12).

Tactile Grasping Feedback Tactile feedback systems use the haptic channel to provide grasping information. The pressure-based device uses openly mounted wires that can create a pressure sensation at the finger tips while the vibration-based system realizes a vibration sensation with small voice-coil motors at the very same location. Similar to the visual feedback, these systems are expected to support the user with grasping judgment and therefore to provide a better interaction. Since these devices use the haptic channel, significant differences depending on the display used, do not have to be expected. The devices are more obtrusive compared to the bare finger tracking, which should be reflected by lower hardware comfort ratings. And finally, the pressure-based system is expected to provide less disturbing feedback than the vibration-based system. For the tactile feedback systems the following effects have to be expected:

Thesis 13 *The pressure-based feedback system supports the user with grasping judgment as well as the vibration-based feedback device and this leads to better interaction capabilities in both conditions.*

Thesis 14 *Since the devices use the haptic channel, no display-related effects can be encountered.*

Thesis 15 *Both feedback devices are obtrusive for the users, while the vibrational feedback is more disturbing during interaction than the pressure-based feedback.*

Similar to the visual feedback, task performance is hardly improved by tactile feedback; only a tendency can be encountered for the pressure-based system (*PTF* vs. *FT*: $t = -1.81, p = 0.083$, $df = 23$). Subjectively, both systems support the users during interaction (*PTF* vs. *FT*: $t = 2.07$, $p = 0.0496$, $df = 23$ and *VTF* vs. *FT*: $t = 2.59$, $p = 0.016$, $df = 23$) (not significant for PTF when adjusted significance-levels are used) and — indicated by intra-subject task performance standard deviations — robustness is improved by them as well, significantly with the pressure-based feedback (*PTF* vs. *FT*: $t = -2.45$, $p = 0.022$, $df = 23$ and *VTF* vs. *FT*: $t = -1.7$, $p = 0.10$, $df = 23$). In general the grasping feedback systems do not improve interaction as much as the grasping support provided by the pinch-sensitive system (Figure 5.20).

Looking at the results of the questionnaire, a strong support of the ability to judge grasps can be seen for both systems (*Q5.PTF* vs. *Q5.FT*: $t = 6.76$, $p < 0.001$, $df = 23$ and *Q5.VTF* vs. *Q5.FT*: $t = 7.22$, $p < 0.001$, $df = 23$). A difference between the tactile devices cannot be stated. Compared to the visual feedback, the users rated grasping judgment ability higher for both tactile devices in the CAVE (*Q5.PTF.C* vs. *Q5.VF.C*: $t = 4.93$, $p < 0.001$, $df = 23$ and *Q5.VTF* vs. *Q5.FT*: $t = 3.08$, $p = 0.011$, $df = 23$). The subjective rating of the actual grasping is hardly improved by tactile grasping feedback, only the vibrational feedback device showed a remarkable improvement (*Q4.VTF* vs. *Q4.FT*: $t = 2.23$, $p = 0.036$, $df = 23$) (not significant with adjusted significant levels). Again this is due to the fact that grasping feedback was provided instead of collision feedback.

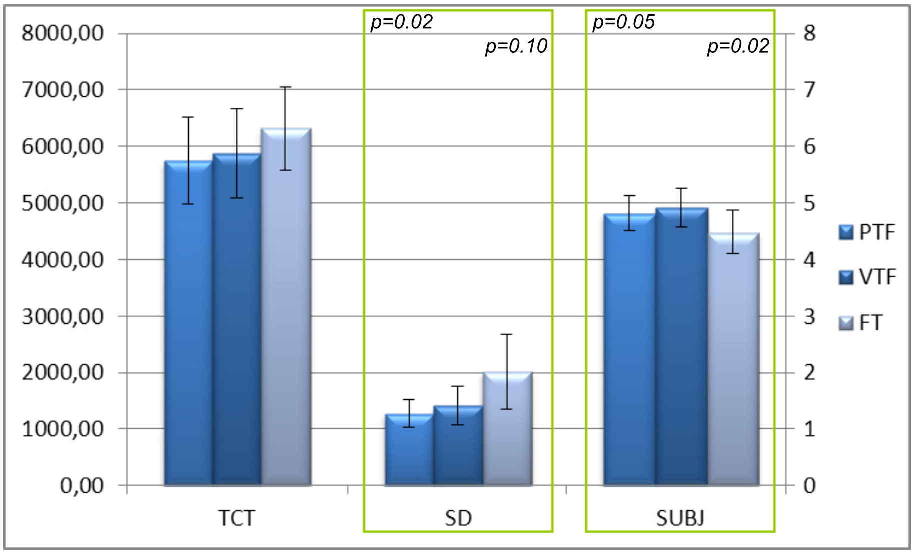

Figure 5.20: Task performance in ms (TCT), intra-subjective standard deviations in ms (SD) and subjective interaction judgment (SUBJ) (0=very bad,...,7=very good) for the pressure-based tactile feedback (PTF), the vibrational device (VTF) and the bare finger tracking (FT). Please note, that the subjective measure refers to the right-hand ordinate. Green frames indicate significant effects. The left significance value within the frames refers to the difference between PTF and FT, the right value refers the difference of VTF and FT condition.

Concerning the comfort ratings, differences between the feedback devices can be identified. The pressure-based feedback was expected to earn lower hardware obtrusiveness ratings because of the openly mounted wires. This prejudice was not confirmed by the study, considering no differences can be found when Q1-ratings of PTF and VTF are compared (*Q1.VTF* vs. *Q1.PTF*: $t = 1.45$, $p = 0.16$, $df = 23$). In contrast, the pressure-based system was expected to provide less obtrusive feedback since the pressure is a more comfortable sensation than the vibration. This present comfort is clearly reflected by the Q8-ratings given by the users (*Q8.VTF* vs. *Q8.PTF*: $t = -4.7$, $p < 0.001$, $df = 23$). Conversely, this also leads to a higher noticeability of the feedback when the vibration-based feedback device is used (*Q2.VTF* vs. *Q2.PTF*: $t = 6.94$, $p < 0.001$, $df = 23$).

In summary, both tactile feedback systems support the user with grasping and provide subjectively better interaction independent of the used display type. Hence theses 13 and 14 are proved by the user study. While the vibration-based system provides stronger feedback, the pressure-based system is more comfortable which confirms the expectations of thesis 15.

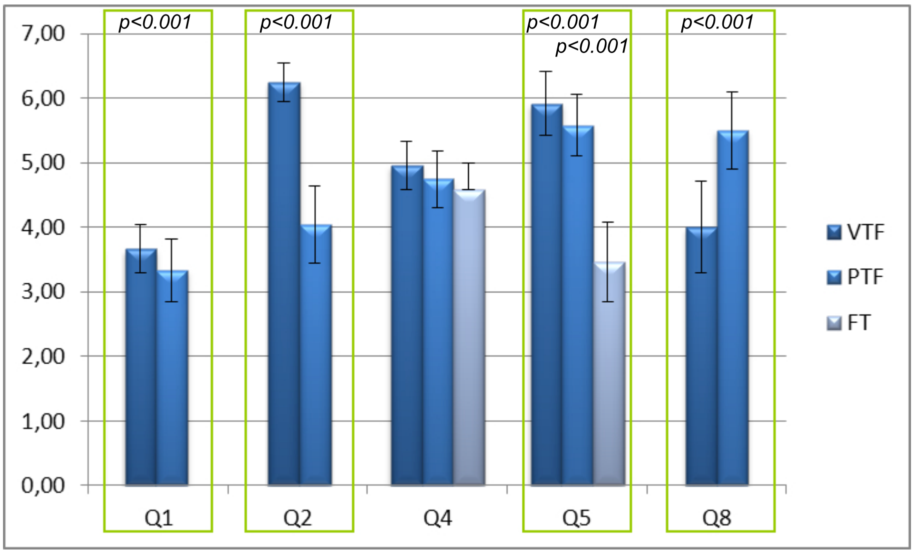

Figure 5.21: Subjective ratings of the questionnaire (0=very bad,...,7=very good) for the pressure-based (PTF) and the vibrational feedback device (VTF) and the bare finger tracking (FT), concerning the hand device obtrusiveness (Q1), the strength of the feedback (Q2), the grasping capabilities (Q4), grasping judgment (Q5) and the obtrusiveness of the feedback (Q8). Green frames indicate significant effects. The left significance value within the frames refers to the difference between PTF and FT, the right value refers the difference of VTF and FT condition.

5.2.3 Conclusion

Direct finger-based interaction enables the users to directly grasp and manipulate virtual objects in a realistic way. While this realism enables functional assessments of virtual car interiors it is indisputable that abstract indirect metaphors provide a more robust way of interaction.

With the extensive user study it was proven that a controller-based metaphor is more robust and efficient than direct finger-based interaction. It also was possible to verify that indirect interaction is less realistic and that it is not sufficient for functional assessments of virtual car interiors in immersive environments. The pinch-sensitive system that was developed to support the user with direct grasping is able to increase the grasping capabilities of the users and it improves interaction reliability and efficiency without loosing the realistic character of the finger-based technique.

Another way to tackle the challenges of direct grasping was suggested with providing grasping feedback. The study showed that grasping feedback does not support the user to the same amount as it is possible with the pinch-sensitive device. While visual feedback is sufficient in HMD-based environments it cannot be used in projection-based environments. A feedback method that supports the user disregarded the display system used, is tactile feedback. Both systems — the pressure-based and the vibrational feedback — work comparably well, but the vibrational feedback device is more disturbing during interaction. Concerning the display devices, no crucial influence on user interaction could be encountered. Both systems can be used for functional car judgments comparably well.

5.3 Visibility of the Virtual Hand

The expert review which was performed in the CAVE revealed no clear preference whether the users' virtual hands should be visualized or not. Half of the experts claimed that a hand representation helps to judge the hands' location with respect to the virtual objects which is a prerequisite for reliable grasping. The other half stated that they were disturbed by the virtual hand models since they tempted them to look at their hands instead of the objects (cf. Section 5.1.2).

Often an offset is applied to the virtual hand to avoid occlusion problems. This is not appropriate for the realistic interaction focused at with this work, since this would falsify the results of the functional validation. Motivated by the statements of the experts, an additional survey reusing the measures and methods of the previous study was conducted to find out if there are objective or subjective influences of the presence of a virtual hand on direct finger-based interaction.

5.3.1 Method

Participants

The same twelve subjects that already participated in the previous study again volunteered for this survey. Between the two studies a break of several days up to one week was introduced.

Apparatus

Of course, a virtual hand model cannot be avoided if a HMD is used, because this type of display completely excludes the real world including the users' real hands. That is why, this study only was conducted in the CAVE-like display system already described previously. For the virtual hand test the basic three-finger optical finger tracking device was used, because of its better usability.

Design

For this study the design of the device influence experiment was reused. The subjects had to perform the same five interaction tasks under two conditions — with a virtual hand representation and without. Again the measured task completion times and subjective judgments were used to compare the usability in each condition. Additionally the subjects were directly asked for their preference and additional comments. In this scenario no grasp recognition improving proxies were used to have a broad range of difficulties in the test to discover the limits of each condition.

Figure 5.22: The two conditions of the study on the influence of the virtual hand representation on the interaction metaphor: hand condition (left) and no-hand condition (right)

As a consequence, the tiny light switch and the thin door handle were hard to grasp for some users.

Procedure

There were two conditions in this study (Figure 5.22). One — the hand condition — had a virtual hand representation registered to the real hands of the user. Although in the projection-based virtual environment the real hand occludes the virtual hand, it was visible for the users due to uncalibrated finger diameters and the simplifications of the hand model. For the other condition — the no-hand condition — the virtual hand geometry was hidden and thus not visible for the users. Nevertheless, the virtual hand model worked as a basis for the grasping heuristics.

First the finger tracking hardware was carefully calibrated to the users' individual hands. After a short familiarization procedure half of the group started the test with a virtual hand representation; the other half started without a virtual hand. The test under the first condition was directly followed by the other condition without removing the finger tracking system in between. Under each condition the subjects had to complete each of the five tasks three times in a row. Task completion times were taken for each trial and after the three trials of each task the participants had to judge their task performance subjectively, as it was done in the previous test. After the subjects performed all trials in all conditions, they were asked for their preference and any further comments.

5.3.2 Results and Discussion

User Preference and User Observation

The preference question shows a clear advance for bare hand interaction. Only one third of the subjects preferred the presence of a virtual hand representation. The main advantage of the absence of a virtual hand was that interaction was perceived as being more direct and natural. The participants mentioned the direct relation between the users' real hand movements and the virtual object reaction as the main reason for rejecting the virtual hand representation. The virtual hand representation was said to be more indirect similar to a mouse cursor. Some subjects claimed that the virtual hand was not perfectly matching the real hand and directs the users' attention to the hands instead of the objects. Since there is a distance between the real hands and the projection screen, this can cause focus shifts that disturb the correct perception of the stereo images.

The advantage of the virtual hand representation on the other hand could be clearly seen when the users interacted with difficult objects — the light switch and the door handle. For most of the users it was hard to grasp these objects without seeing the virtual hand. Five subjects were not able to interact with the light switch at all and three subjects had the same problem with the door in contrast to one and no subject respectively when the virtual hand was visible. The virtual hand representation provides more feedback on were the users' hands are located with respect to the objects and facilitate the interaction with challenging objects.

Task Completion Times

Task completion times are comparable under both conditions for the objects that can be grasped without any problems (sun shield, mirror, bottle — Figure 5.23). The inter- and inner-subject variances are low and similar for both conditions. The sun shield task could be completed significantly faster without a virtual hand representation ($t = 2.8$, $p = 0.02$, $df = 11$) but with a minor difference of half a second.

The problems some users had with the light switch and the door were directly reflected by the task performance, leading to significantly higher task completion times in the no-hand condition ($t = -2.2$, $p = 0.05$, $df = 10$ and $t = -2.2$, $p = 0.05$, $df = 11$ respectively). Variances are very high for these objects in both conditions. In seven cases subjects were not able to interact with one of these objects in the no-hand condition. For calculating the mean values of task performance in this case, the highest TCT the participant achieved for this particular object in hand condition was used as the TCT for the no-hand condition.

Subjective Measures

Subjective measures reflect the users' preference for the no hand condition resulting in better judgments for easy-to-use objects (Figure 5.24). For the door and the light switch the grasping problems lead to better judgments for the hand condition. Concerning subjective measures, the T-test showed a significant preference for the hand condition in the door task ($t = 3.6$, $p = 0.004$, $df = 11$) and a significant better value for the mirror task ($t = -4$, $p = 0.002$, $df = 11$) when it was performed without a visible hand.

In general it can be stated that users do not prefer having a virtual hand representation. If grasping is robust they do not need it to judge the location of their hands with respect to the virtual objects. In fact having no virtual hand leads to a higher perceived realism and naturalness

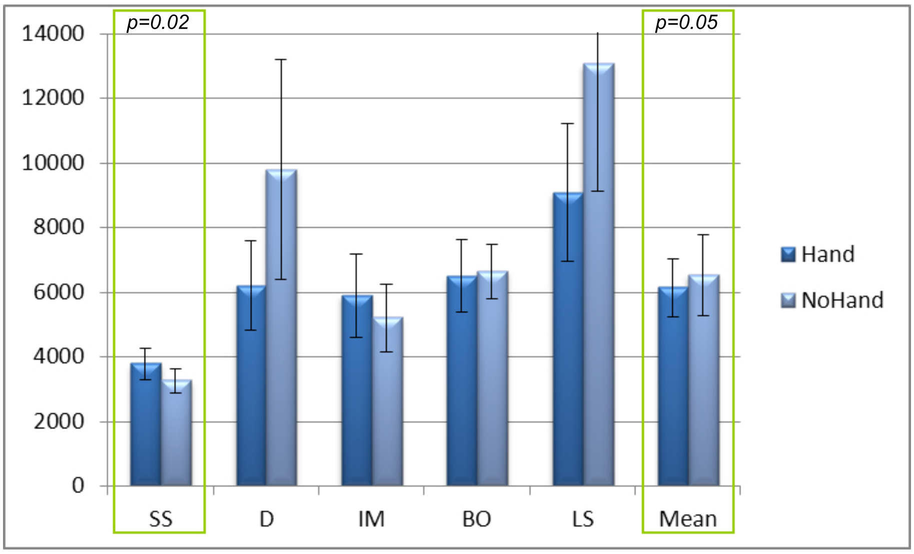

Figure 5.23: Per-object (sun shield (SS), door (D), interior mirror (IM), bottle (BO) and light switch (LS)) and average (Mean) task completion times in ms for the hand vs. the no hand trial. Green frames indicate significant differences.

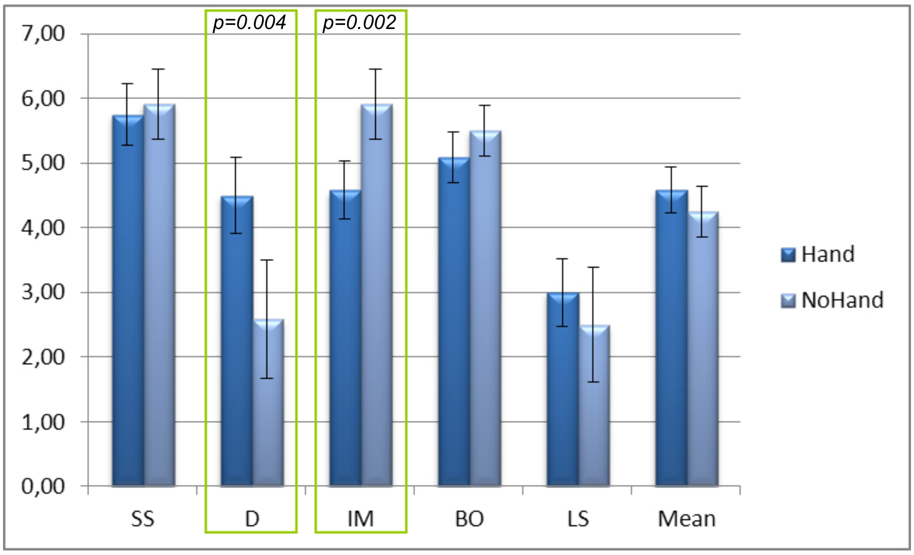

Figure 5.24: Per-object (sun shield (SS), door (D), interior mirror (IM), bottle (BO) and light switch (LS)) and average subjective interaction ratings (0=very bad,...,7=very good) for the hand vs. the no hand trial. Green frames indicate significant differences.

of the interaction metaphor. In difficult cases a virtual hand can help users to understand why an object cannot be properly grasped. Of course, without a virtual hand, the projection-based virtual environments have to have precisely calibrated screen setups and tracking systems.

5.3.3 Conclusion

To sum up, having no virtual hand representation could be preferred for direct finger-based interaction when grasping is robust and reliable enough. In these situations, interaction realism seems to be supported by hiding the virtual hand geometry. Only in situations where the users are not able to immediately grasp an object, the presence of the hand is necessary for grasping the objects. Especially when the users are supported by the enhanced finger tracking hardware, proposed in this work, a virtual hand representation should be avoidable in most of the cases, leading to a more realistic virtual experience. However this remains unproven and should be further addressed in future work.

5.4 Conclusion

In general the evaluation indicated that the direct finger-based interaction is a useful interaction concept for functional car assessments in virtual environments. Unsupported interaction cannot keep up with the efficiency and robustness of controller-based interaction. The grasping capabilities of the users along with task performance and robustness can be increased with enhanced hand tracking hardware that detects finger pinches or provides grasping feedback. Visual feedback can also improve the interaction capabilities of the users. With a combination of the pinch-sensitive mechanism with grasping feedback — tactile or visual — the direct finger-based interaction probably could become almost as robust and efficient as controller-based interaction and could even be used without a virtual hand representation.

Chapter 6

Conclusions

In conclusion, the research questions that inspired this work are reviewed and remaining aspects and open questions are discussed afterwards in a future work section. The closing remarks summarize the findings and name the limits of this approach.

6.1 Contributions

Throughout this thesis, the development of a robust, efficient and plausible approach to direct finger-based interaction is described in great detail. This interaction technique was realized to be all-encompassing enough to enable manipulation of the large variety of constrained and unconstrained objects typically found in car interior evaluation scenarios. Furthermore, the thesis addresses the challenges of direct grasping in virtual environments by introducing the concept of Normal Proxies, it enhances the grasping heuristics with information from a pinch-sensitive finger tracking device and it employs tactile grasping feedback. A comprehensive evaluation compares the new interaction concept and its variations with existing techniques and reveals the inherent potential of the hardware enhancements. These contributions directly address the main research questions posed in the beginning of the thesis:

What are the algorithmical and technical improvements for direct finger-based interaction that make the usability acceptable for a variety of automotive evaluation tasks?

Realistic and robust interaction by grasping heuristics and pseudophysical object behavior At present only few immersive interactive virtual assessments have made their way into automotive development processes. Those that have usually rely on indirect interaction metaphors that are very robust and reliable but not realistic enough to allow for an analysis of the corresponding real-world interaction. Wherever human-car interaction is the focus and a realistic experience is required, direct manual manipulation is the only choice. The developed direct finger-based interaction combines powerful grasping heuristics with a pseudophysical calculation of object behavior.

In regards to grasping heuristics, human grasping behavior in car interiors was carefully analyzed and the documented grasps were classified consistent with common grasp classifications.

Simple rules of grasping were derived that are easy to compute and simultaneously represent the large variance of human grasping behavior. The resulting heuristics provide realistic grasping that is both robust and reliable.

The pseudophysical calculations of object behavior do not simulate physical processes but geometrically calculate a plausible object motion as a reaction to the input of the grasping hand. Therefore the motion of the hands and fingers is analyzed and adequately applied on the virtual object. Objects in a car interior usually are not freely movable, and in fact, the motion of the majority of the objects is restricted in several ways. A thorough analysis of the objects in the car interior revealed different kinds of constraints that influence object behavior. The pseudophysical object reaction was enhanced to cover objects which motion is restricted by mounting, functional and location constraints. Moreover, another specific characteristic of human-car interaction was addressed by direct finger-based interaction; namely, push interaction, in which objects are pushed by single fingers or the whole hand. With this specific interaction and the different types of constraints, the entire range of human-object interaction found in a car interior could be addressed using direct finger-based interaction.

The direct finger-based interaction technique enables realistic and intuitive manipulation of arbitrary virtual objects. However, direct grasping in virtual environments is challenging for some users. This is caused by missing haptic feedback, stereoscopy issues due to focus and convergence mismatch and tracking jitter, all of which impede the grasping process in projection-based and other virtual environments. This is true especially when considering thin and tiny objects since they are sometimes hard to pick up and require extremely precise selection.

Better grasping of thin and tiny objects by pinch-sensitivity One possibility to address these challenges arose from the development of a pinch-sensitive optical finger tracking system. This system is able to detect pinches of the users' real fingers. The users tend to pinch their fingers when they grasp objects, particularly if they are small or thin. Since the natural act of pinching cannot be avoided, it was inevitably put to use. When a pinch of the thumb with another finger is detected and at least one of the fingers is colliding with a virtual object, it is obvious that the user intends to grasp the object. With this enhancement of the grasping rules, direct finger-based grasping in virtual environments was improved significantly, particularly for the troublesome thin and tiny objects.

Tackling normal problems with Normal Proxies The grasping rules are based on the face normals which are needed to simulate friction in a pseudophysical way. However, the use of the original face normals can lead to unsatisfactory grasping reliability within various scenarios. One issue is the occurrence of incorrect surface normals which are still common in everyday industrial processes. Another issue relates to box-shaped objects or tiny objects. Normal Proxies address all these face normal issues by introducing an additional normal for object regions that is only used for grasping calculations. It is through this concept that grasping robustness was considerably improved.

Direct finger-based interaction as a combination of reliable grasping heuristics with plausible pseudophysical object behavior enables functional assessments of virtual car interiors by providing robust and realistic interactions. With the integration of a pinch-sensitive device and the concept of normal proxies, interaction reliability was further improved making the usability of this interaction technique acceptable for everyday industrial applications.

How to measure the effectiveness and usability of direct finger-based interaction in comparison to other techniques?

Using two different methods, various aspects of direct finger-based interaction were evaluated. An initial expert review analyzed the overall usability of this approach and its applicability in industrial applications. It indicated that users who have some experience with stereoscopic technology became quickly proficient with the interaction technique. Automotive experts could easily validate various functional aspects of a car, such as the accessibility and usability of car interfaces, visibility aspects and object clearances.

Additionally a comprehensive user study used quantitative and qualitative measures to analyze the interaction concept. As a result, the study combined task completion times with subjective judgments and preference questionnaires for a broad understanding of all aspects of interaction. The results of the user study clearly showed that direct finger-based interaction is preferred over indirect interaction for the assessments of various functional characteristics of a car interior. While controller-based interaction is clearly faster and more robust, the abstract character of indirect metaphors leads to a loss of realism and therefore impairs the judgment of the experts. The deficits of direct methods concerning performance and robustness can be almost compensated by combining pinch-recognition with a precise finger tracking system.

Which types of grasping feedback are helpful and preferred for performing finger-based interaction tasks?

One of the challenges for direct grasping is the absence of any kind of haptic feedback from the virtual environment. This work addresses this deficit with another enhancement of the hand-tracking hardware. This extended hand-tracking hardware presented tactile grasping feedback to the users to better support them during interaction.

Two approaches towards tactile feedback were also realized within this thesis. An innovative approach using wires of shape memory alloys enabled the development of light-weight and unobtrusive actuators that provide tactile grasping feedback at the finger tips. The SMA wires shorten when a current is applied to them. This shortening results in a pressure sensation at the user's finger tips. An iterative design of the feedback actuators and the development of a micro-controller to control the feedback device lead to a prototype that was piloted as a commercially available device.

As an alternative, small voice-coil motors were integrated into the finger thimbles of a finger tracking device. They work as actuators for regular vibrational feedback. In contrast to available devices, here the feedback again takes place were the users are expecting it — at the finger tips. As an additional reference, visual grasping feedback was integrated into the interaction framework. Therefore, grasping-related fingers of the virtual hand representations were colored orange during grasping.

The grasping feedback approaches were evaluated within the user study as well. The results revealed that grasping feedback is a requirement to judge grasp status. It is not sufficient to simply have an object following the user's hand motion once it is grasped. While visual feedback alone is sufficient for HMD-applications, tactile feedback significantly improves interaction independent of the display system. Vibrational feedback is considerably stronger than pressure-based sensations but can quickly become annoying.

6.2 Future Work

Direct finger-based interaction makes it possible to realistically experience a virtual car interior and to assess functional aspects during car development. Considering the aforementioned enhancements to both (the interaction metaphor and the hand-tracking hardware), it is possible to interact with virtual objects in a realistic and reliable manner. Nevertheless, the following section contains suggestions for research considerations that are worthwhile to pursue.

Interaction-Supporting Hand-tracking Hardware This work proved that interaction is significantly supported by the proposed enhancements to the finger tracking hardware. The pinch-sensitive mechanism improved interaction capabilities as well as tactile feedback presented at the finger tips. A combination of both approaches has the potential to further improve direct interaction and increase its acceptance. A hybrid device would provide robust grasping by accurate pinch detection and clearly communicated interaction states as a result of the feedback. With this device, the task performance of direct finger-based interaction would probably further improve when compared to indirect interaction and reality. However, it remains challenging to integrate both mechanisms into the thimbles of the finger tracking device.

Throughout this thesis, the feedback is used to report the grasp status to the users. This grasping feedback addresses the challenges of users in judging if they succeeded in grasping and this work provides strong evidence that their interaction capabilities have been considerably improved. An alternative to grasping feedback would be to communicate finger collisions with the tactile feedback systems. This collision feedback would address the challenges of users in judging the location of their fingers with respect to the virtual objects. Further evaluation of the challenges and shortcomings of both approaches remains necessary.

Both tactile feedback systems presented in this work could be further improved to provide tactile patterns. This can be achieved by separately addressing the different wires of the pressure-based system or by alternating the frequency of the vibrational actuators of the vibration-based feedback device. With these tactile patterns, different states throughout the grasping process could be communicated which could further improve the feedback. The very same effect could also be achieved through a combination of different feedback methods.

Until unobtrusive, precise and crisp full finger haptic feedback becomes available, the finger-based interfaces presented here provide a balanced and robust compromise for the realistic assessment of various functionalities of newly designed objects in virtual environments.

Physical Simulation-based Interaction Metaphor The range of pseudophysical behavior of objects in automotive applications is well covered by the developments in this thesis. Further physical effects, such as realistic object-object interaction, the influence of gravity and inertia or deformable hand models for more realistic hand behavior can only be provided by physical simulations. However, these solutions are still not robust enough for the challenging tasks of industrial applications. We expect that the gaming market will act as a driving force and further increases in readily available computational resources will enable real physical interaction metaphors in the not so distant future — even for the complex virtual scenarios dealt with here.

Pseudohaptic Effects The pseudophysical reaction of objects could be further extended to provide several pseudohaptic effects [LCR$^+$00]. In fact, physical behavior is expected by the users of direct finger-based interaction, which was revealed by the expert review. Different acceleration and inertia parameters, resulting from individual object characteristics, could be applied to the object reaction such that the object immediately follows the users' hand motion to simulate light-weight objects while heavy objects would exhibit more inertia. However, this would lead to diverging locations of the user's hands and their virtual representation. Here concepts are necessary in order to convey how this divergence could be minimized while employing pseudohaptic effects. Moreover, it needs to be researched to find out if such pseudohaptic effects are preferred by the users over directly co-located real and virtual hands without individual object reaction.

Video-based Finger Tracking Several research approaches have been made to overcome the necessity to wear finger tracking hardware for the transfer of the users' hands into the virtual environment. The recent advances induced by increasing requirements of the gaming market make it probable that unobtrusive finger tracking without any finger-worn hardware will be available in the future. This thesis was able to show with a user study, that users prefer to omit the virtual hand representation when they are able to robustly grasp the virtual objects. It is obvious that they also would appreciate the absence of finger-worn hardware. Another open question is, to which amount the perceived realism of direct finger-based interaction is influenced by the absence of a virtual hand representation and to which amount it will be influenced by the disappearance of the hardware. However, the presence of the finger tracking hardware turned out to be a chance for interaction support during this present work. Of course this possibility becomes nonexistent if video-based finger tracking were used.

6.3 Closing Remarks

This thesis addresses a broad range of aspects of interaction in automotive industrial scenarios. Considering the current situation of the automotive industry which is driven by severe competition with decreasing margins and increasing product complexity and requirements, it is obviously inevitable to rely more and more on virtual car data and on virtual assessments. The majority of necessary assessments during product development do not need sophisticated, immersive VR applications. For these it is sufficient enough to provide the possibility to visually analyze the realistically presented virtual car data.

Nevertheless, a further pervasion of automotive product development processes with virtual analyses needs the possibility to virtually experience the *functionality* of a car concept. The functional and emotional experience of car concepts can only be realized with virtual cars if they become interactive. Therefore realistic or at least plausible interaction metaphors in immersive environments are essential. At the same time these interaction metaphors have to be robust and reliable to be acceptable for real world industrial processes and applications. This thesis prepares the ground for virtual assessments concerning the car's functionality and the human-car interplay by addressing both interaction goals. While this research considerably advances the state of the art in simulating direct finger-based interaction in virtual environments, there is still a long way ahead until these interactions are as convincing as interactions with the real world.

Bibliography

[AHS03] Albrecht, Irene, Joerg Haber and Hans-Peter Seidel: *Construction and animation of anatomically based human hand models.* In *Proceedings of the ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, SCA, pages 89–109, Aire la Ville, Switzerland, 2003.

[AMK^{+}09] Aoki, Takafumi, Hironori Mitake, Danial Keoki, Shoichi Hasegawa and Makoto Sato: *Wearable haptic device to present contact sensation based on cutaneous sensation using thin wire.* In *Proceedings of ACM International Conference on Advances in Computer Entertainment Technology*, ACE'09, pages 115–122, Athens, Greece, 2009.

[AMR^{+}07] Allard, Jeremie, Clement Menier, Bruno Raffin, Edmond Boyer and Francois Faure: *Grimage: Markerless 3D Interactions.* In *Siggraph 2007 (Emerging Technologies)*, San Diego (CA), USA, 2007. ACM.

[Bar97] Baraff, David: *An Introduction to Physically Based Modeling: Rigid Body Simulation I - Unconstrained Rigid Body Dynamics.* In *In SIGGRAPH '97 Course Notes*, 1997.

[BB04] Brewster, Stephen A. and Lorna M. Brown: *Tactons: Structured Tactile Messages for Non-Visual Information Display.* In *Proceedings of the 5th Australasian User Interfaces Conference*, AUIC, pages 15–23, Dunedin, New Zealand, 2004.

[BDF^{+}01] Bowman, D.A., A. Datey, U. Farooq, Y.S. Ryu and O. Vasnaik: *Empirical Comparisons of Virtual Environment Displays (TR-01-19).* Technical Report, Department of Comuter Science, Virginia Tech, 2001.

[BFH02] Borst, Ch., M. Fischer and G. Hirzinger: *Calculating Hand Configurations for Precision and Pinch Grasps.* In *Proceedings of the 2002 IEEE/RSI/GI International Conference on Intelligent Robots and Systems*, pages 1553–1559, Lausanne, Switzerland, 2002.

[BH97] Bowman, Doug A. and Larry F. Hodges: *An Evaluation of Techniques for Grabbing and Manipulating Remote Objects in Immersive Virtual Environments.* In *Proceedings of the Symposium on Interactive 3D Graphics*, I3D, page 35ff., New York (NY), USA, 1997.

[BHB^{+}95] Barfield, Woodrow, Claudia M. Hendrix, Ove Bjorneseth, Kurt A. Kaczmarek and Wouter Lotens: *Comparison of Human Sensory Capabilities with Technical Specifications of Virtual Environment Equip-*

ment. Presence, 4(4):329–356, 1995.

[BI06] Borst, Christoph W. and Arun P. Indugula: *A Spring Model for Whole-Hand Virtual Grasping*. Presence, 15(1):47–61, 2006.

[BIS01] Bouguila, Laroussi, Masahiro Ishii and Makoto Sato: *Scaleable SPIDAR: A Haptic Interface For Human-Scale Virtual Environments*. In *Proceedings of the 1st International Workshop on Haptic Human-Computer Interaction*, pages 182–193, 2001.

[BKHAK04] Benali-Khoudja, Mohamed, Moustapha Hafez, Jean Marc Alex and Abderrahmane Kheddar: *Tactile Interfaces: a State-of-the-Art Survey*. In *Proceedings of International Symposium on Robotics*, pages 721–726, 2004.

[BKLP04] Bowman, Doug A., Ernst Kruijff, Joseph J. LaViola and Ivan Poupyrev: *3D User Interfaces: Theory and Practice*. Addison Wesley Longman Publishing Co., Inc., Redwood City, CA, USA, 2004.

[BRP+06] Burns, Eric, Sharif Razzaque, Abigail Panter, Mary C. Whitton, Matthew R. McCallus and Frederick P. Brooks Jr.: *The Hand Is More Easily Fooled than the Eye: Users Are More Sensitive to Visual Interpenetration than to Visual-Proprioceptive Discrepancy*. Presence, 15(1):1–15, 2006.

[Bur96] Burdea, Grigore C.: *Force and Touch Feedback for Virtual Reality*. John Wiley & Sons, Inc., New York (NY), USA, 1996.

[BWC+02] Bowman, D. A., C. A. Wingrave, J. M. Campbell, V. Q. Ly and C. J. Rhoton: *Novel Uses of Pinch Gloves for Virtual Environment Interaction Techniques*. In *Proceedings of IEEE Virtual Reality Conference*, VR, pages 122–129, 2002.

[CFH97] Cutler, Lawrence D., Bernd Fröhlich and Pat Hanrahan: *Two-Handed Direct Manipulation on the Responsive Workbench*. In *Proceedings of Symposium on Interactive 3D Graphics*, SI3D, pages 107–114, 1997.

[CMH05] Chouvardas, Vasilios G., Amalia N. Miliou and Miltiadis K. Hatalis: *Tactile Displays: a Short Overview and Recent Developments*. In *Proceedings of the 5th International Conference on Technology and Automation*, pages 246–251, 2005.

[CNSD93] Cruz-Neira, Carolina, Daniel J. Sandin and Thomas A. DeFanti: *Surround-Screen Projection-based Virtual Reality: The Design and Implementation of the CAVE*. In *Proceedings of ACM SIGGRAPH*, SIGGRAPH, pages 135–142, 1993.

[DCdlPAB08] Duriez, Christian, Hadrien Courtecuisse, Juan-Pablo de la Plata Alcalde and Pierre-Jean Bensoussan: *Contact Skinning*. In *Proceedings of Eurographics*, EUROGRAPHICS, 2008.

[DJK+06] Demiralp, C., C.D. Jackson, D.B. Karelitz, S. Zhang and D.H. Laid-

LAW: *CAVE and Fishtank Virtual-Reality Displays: A Qualitative and Quantitative Comparison.* IEEE Transactions on Visualization and Computer Graphics (TVCG), 12(3):323–330, 2006.

[dlGPF08] GORCE, MARTIN DE LA, NIKOS PARAGIOS and DAVID J. FLEET: *Model-Based Hand Tracking with Texture, Shading and Self-Occlusions.* In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, CPVR, Anchorage, Alaska, 2008.

[EKK⁺07] ENDO, YUI, SATOSHI KANAI, TAKESHI KISHINAMI, NATSUKI MIYATA and MAKIK KOUCHI: *A Computer-aided Ergonomic Assessment and Product Design System Using Digital Hands.* In *Proceedings of the 1st International Conference on Digital Human Modeling*, ICDHM, pages 833–842, Bejing, China, 2007.

[Esc93] ESCHER, K.: *Two-way Shape Memory Effects for the Construction of Grasping Elements (Die Zweiweg-Formgedaechtniseffekte zur Herstellung von Greifelementen).* In *VDI Fortschrittsberichte Nr. 298*, Duesseldorf, Germany, 1993.

[FTB⁺00] FRÖHLICH, BERND, HENRIK TRAMBEREND, ANDREW BEERS, MANEESH AGRAWALA and DAVID BARAFF: *Physically-Based Manipulation on the Responsive Workbench.* In *Proceedings of the IEEE Virtual Reality Conference*, VR, pages 5–11, Washington DC, USA, 2000.

[Gol03] GOLDSTEIN, E. BRUCE: *Sensation and Perception.* Cengage Learning, Inc., Florence (KY), USA, 2003.

[Hae10] HAENTSCH, FERRY: *Finger Collision Detection for Virtual Environments (Fingerkollisionserkennung fuer virtuelle Umgebungen).* Weimar, Germany, 2010.

[HL92] HUNTER, I. W. and S. LAFONTAINE: *Shape Memory Alloy Fibres Having Rapid Twitch Response*, 1992.

[Hol79] HOLM, SORE: *A Simple Sequentially Rejective Multiple Test Procedure.* In *Scandinavian Journal of Statistics*, pages 65–70, 1979.

[HSA⁺08] HULIN, T., M. SAGARDIA, J. ARTIGAS, S. SCHAETZLE, P. KREMER and C. PREUSCHE: *Human Scale Bimanual Haptic Interface.* In *Proceedings of Enactive*, Pisa, Italy, 2008.

[HUWK08] HOLZ, DANIEL, SEBASTIAN ULLRICH, MARC WOLTER and TORSTEN KUHLEN: *Multi-Contact Grasp Interaction for Virtual Environments.* Journal of Virtual Reality and Broadcasting, 5(7), 2008.

[IF06] INABA, GO and KINYA FUJITA: *A Pseudo-Force-Feedback Device by Fingertip Tightening for Multi-Finger Object Manipulation.* In *Proceedings of EuroHaptics*, Paris, France, 2006.

[JF11] JACOBS, JAN and BERND FROEHLICH: *A Soft Hand Model for Physically-based Manipulation of Virtual Objects.* In *Proceedings of the IEEE Virtual Reality Conference*, VR, pages 11–18, Singapore, Singapore, 2011.

[JNL04] JONES, LYNETTE A., MEALANI NAKAMURA and BRETT LOCKYER: *Development of a Tactile Vest.* In *Proceedings of IEEE Haptics*, pages 82–89, Chicago (IL), USA, 2004.

[KH95] KONTARINIS, DIMITRIOS A. and ROBERT D. HOWE: *Tactile Display of High-Frequency Information in Teleoperation and Virtual Environments.* Presence, 4(4):387–402, 1995.

[KHW95] KESSLER, G. DREW, LARRY F. HODGES and NEFF WALKER: *Evaluation of the CyberGlove as a Whole-Hand Input Device.* ACM Transactions on Computer-Human Interaction, 2:263–283, 1995.

[KI91] KANG, SING BING and KATSUSHI IKEUCHI: *A Framework for Recognizing Grasps (CMU-RI-TR-91-24).* Technical Report, Robotics Institute, Carnegie Mellon University, 1991.

[Kje01] KJELDSKOV, J.: *Combining Interaction Techniques and Display Types for Virtual Reality.* In *Proceedings of the Annual Conference of the Australian CHISIG*, OzCHI, pages 77–83, Churchland, AUS, 2001.

[KKHS04] KAMMERMEIER, PETER, ALEXANDER KRON, JENS HOOGEN and GÜNTHER SCHMIDT: *Display of Holistic Haptic Sensations by Combined Tactile and Kinesthetic Feedback.* Presence, 13(1):1–15, 2004.

[KP03] KRY, PAUL G. and DINESH K. PAI: *Interaction Capture and Synthesis.* In *Proceedings of the ACM SIGGRAPH*, SIGGRAPH, pages 872–880, New York, NY, 2003.

[LCR$^+$00] LÉCUYER, ANATOLE, SABINE COQUILLART, INRIA ROCQUENCOURT, ABDERRAHMANE KHEDDAR, PHILIPPE COIFFET and PAUL RICHARD: *Pseudo-Haptic Feedback: Can Isometric Input Devices Simulate Force Feedback?* In *Proceedings of IEEE Virtual Reality Conference*, VR, pages 83–90, Washington DC, USA, 2000.

[LG94] LIANG, JIANDONG and MARK GREEN: *JDCAD: A highly interactive 3D modeling system.* Computers & Graphics, 18(4):499–506, 1994.

[Liu09] LIU, C. KAREN: *Dextrous Manipulation from a Grasping Pose.* In *Proceedings of the ACM SIGGRAPH*, SIGGRAPH, New York, NY, 2009.

[LZ99] LAVIOLA, JOSEPH and ROBERT C. ZELEZNIK: *Flex And Pinch: A Case Study Of Whole Hand Input Design For Virtual Environment Interaction.* In *Proceedings of the International Conference on Computer Graphics and Imaging*, pages 157–168, 1999.

[MA03] MASCARO, STEPHEN A. and H. HARRY ASADA: *Wet Shape Memory Alloy Actuators for Active Vasculated Robotic Flesh.* In *Proceedings of IEEE International Conference on Robotics and Automation*, ICRA, pages 282–287, Taipeh, Taiwan, 2003.

[MF10] MOEHRING, MATHIAS and BERND FROEHLICH: *Enabling functional valida-*

tion of virtual cars through Natural Interaction metaphors. In *Proceedings of IEEE Virtual Reality Conference*, VR, pages 27–34, Waltham (MA), USA, 2010.

[MF11a] MOEHRING, MATHIAS and BERND FROEHLICH: *Effective Manipulation of Virtual Objects Within Arm's Reach.* In *Proceedings of IEEE Virtual Reality Conference*, VR, pages 131–138, Singapore, Singapore, 2011.

[MF11b] MOEHRING, MATHIAS and BERND FROEHLICH: *Natural Interaction Metaphors for Functional Validations of Virtual Car Models.* IEEE Transactions on Visualization and Computer Graphics (TVCG), 17(9):1195–1208, 2011.

[MPT99] MCNEELY, WILLIAM A., KEVIN D. PUTERBAUGH and JAMES J. TROY: *Six Degree-of-Freedom Haptic Rendering Using Voxel Sampling.* In *Proceedings of ACM SIGGRAPH*, pages 401–408, 1999.

[MS94] MASSIE, THOMAS H. and J.K. SALISBURY: *The PHANTOM Haptic Interface: A Device for Probing Virtual Objects.* In *Proceedings of the ASME Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, Chicago (IL), USA, 1994.

[PBWI96] POUPYREV, IVAN, MARK BILLINGHURST, SUZANNE WEGHORST and TADAO ICHIKAWA: *The Go-Go Interaction Technique: Non-Linear Mapping for Direct Manipulation in VR.* In *Proceedings of the 9th ACM Symposium on User Interface Software and Technology*, UIST, pages 79–80, 1996.

[PRS02] PREECE, J.J., Y. ROGERS and H. SHARP: *Interaction Design: beyond human computer interaction.* John Wiley & Sons, Inc., New York (NY), USA, 2002.

[PSL$^+$07] PABON, SILVIA, EDOARDO SOTGIU, ROSARIO LEONARDI, CRISTINA BRANCOLINI, OTNIEL PORTILLO-RODRIGUEZ, A. FRISOLI and MASSIMO BERGAMASCO: *A data-glove with vibro-tactile stimulators for virtual social interaction and rehabilitation.* In *10th annual International Workshop on Presence*, Barcelona, Spain, 2007.

[PWBI98] POUPYREV, IVAN, SUZANNE WEGHORST, MARK BILLINGHURST and TADAO ICHIKAWA: *Egocentric Object Manipulation in Virtual Environments: Evaluation of Interaction Techniques.* Computer Graphics Forum, 17(3):41–52, 1998.

[QIHM97] QI, W., R.M. TAYLOR II, C.G. HALEY and J.-B. MARTENS: *A Comparison of Immersive HMD, Fish Tank VR and Fish Tank with Haptics Display for Volume Visualization.* In *Proceedings of 3rd Symposium on Applied Perception in Graphics and Visualization*, pages 107–114, Boston (MA), USA, 1997.

[RBH$^+$95] REZZONICO, SERGE, RONAN BOULIC, ZHIYONG HUANG, NADIA MAGNENAT THALMANN and DANIEL THALMANN: *Consistent Grasping Interactions with Virtual Actors.* In *Proceedings of the 2nd EUROGRAPHICS Workshop on Virtual Environments*, EGVE, pages 709–734, Monte Carlo, Monaco, 1995.

[RHSM05] REGENBRECHT, HOLGER, JÖRG HAUBER, RALPH SCHÖNFELDER and ANDREAS MAEGERLEIN: *Virtual Reality Aided Assembly with Directional Vibro-*

Tactile Feedback. In *Proceedings of the 3rd International Conference on Computer Graphics and Interactive Techniques in Australasia and Southeast Asia*, GRAPHITE, pages 381–387, Dunedin, New Zealand, 2005.

[RHWF06] RIEGE, KAI, THORSTEN HOLTKAMPER, GEROLD WESCHE and BERND FROEHLICH: *The Bent Pick Ray: An Extended Pointing Technique for Multi-User Interaction*. In *Proceedings of Symposium on 3D User Interfaces*, 3DUI, pages 62–65, Washington (DC), USA, 2006.

[RIKA08] RIES, B., V. INTERRANTE, M. KAEDING and L. ANDERSEN: *The Effect of Self-Embodiment on Distance Perception in Immersive Virtual Environments*. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology*, VRST, pages 167—170, Bordeaux, France, 2008.

[RK04] RYU, JONGHYUN and GERARD JOUNGHYUN KIM: *Using a Vibro-Tactile Display for Enhanced Collision Perception and Presence*. In *Proceedings of ACM Symposium on Virtual Reality Software and Technology*, VRST, pages 89–96, Hong Kong, China, 2004.

[SCP95] STOAKLEY, RICHARD, MATTHEW CONWAY and RANDY F. PAUSCH: *Virtual Reality on a WIM: Interactive Worlds in Miniature*. In *Proceedings of SIGCHI Conference on Human Factors in Computing Systems*, CHI, pages 265–272, 1995.

[SF08] SALZMANN, HOLGER and BERND FROEHLICH: *The Two-User Seating Buck: Enabling Face-to-Face Discussions of Novel Car Interface Concepts*. In *Proceedings of IEEE Virtual Reality Conference*, VR, pages 75–82, Reno (NV), USA, 2008.

[Shi93] SHIMOGA, KARUN B.: *A Survey of Perceptual Feedback Issues in Dexterous Telemanipulation: Part II. Finger Touch Feedback*. In *Proceedings of IEEE Virtual Reality Conference*, VR, pages 271–279, 1993.

[Sho85] SHOEMAKE, KEN: *Animating rotation with quaternion curve*. In *Proceedings of the 12th annual conference on Computer Graphics and Interactive Techniques*, SIGGRAPH'85, New York (NY), USA, 1985.

[SMF07] SCHEIBE, ROBERT, MATHIAS MOEHRING and BERND FROEHLICH: *Tactile Feedback at the Finger Tips for Improved Direct Interaction in Immersive Environments*. In *Proceedings of the IEEE Symposium on 3D User Interfaces*, 3DUI, pages 123–129, Charlotte (NC), USA, 2007.

[SRH04] STEINICKE, FRANK, TIMO ROPINSKI and KLAUS HINRICHS: *Object Selection in Virtual Environments with an Improved Virtual Pointer Metaphor*. In *Proceedings of International Conference on Computer Vision and Graphics*, ICCVG, pages 320–326, 2004.

[ST94] SANSO, R. M. and DANIEL THALMANN: *A Hand Control and Automatic Grasping System for Synthetic Actors*. Computer Graphics Forum, 13(3):167–177, 1994.

[SZS95] SALISBURY, ZILLES, C. B. ZILLES and J. K. SALISBURY: *A Constraint-based God-object Method For Haptic Display.* In *Proceedings of the International Conference on Intelligent Robots and Systems*, IROS, pages 146–151, 1995.

[US00] ULLMANN, THOMAS and JOERG SAUER: *Intuitive Virtual Grasping for Non Haptic Environments.* Proceedings of Pacific Conference on Computer Graphics and Applications, page 373ff, 2000.

[vdPRHP99] POL, ROGIER VAN DE, WILLIAM RIBARSKY, LARRY HODGES and FRITS POST: *Interaction in Semi-Immersive Large Display Environments.* In *Proceedings of the EUROGRAPHICS Workshop on Virtual Environments*, EGVE, pages 157–168, 1999.

[VPHS05] VELAZQUEZ, RAMIRO, EDWIGE PISSALOUX, MOUSTAPHA HAFEZ and JEROME SZEWCZYK: *A Low-Cost Highly-Portable Tactile Display Based on Shape Memory Alloy Micro-Actuators.* In *Proceedings of the 2005 IEEE International Conference on Virtual Environments, Human Computer Interfaces and Measurement Systems*, VECIMS'05, page 6ff., 2005.

[War90] WARE, COLIN: *Using Hand Position for Virtual Object Placement.* The Visual Computer, 6(5):245–253, 1990.

[Web10a] WEBSITE: *A.R.T. GmbH*, September 2010. `http://www.ar-tracking.com`.

[Web10b] WEBSITE: *CyberGlove Systems*, September 2010. `http://www.cyberglovesystems.com`.

[Web11a] WEBSITE: *Fraunhofer IGD*, May 2011. `http://www.igd.fraunhofer.de`.

[Web11b] WEBSITE: *GrImage*, May 2011. `http://grimage.inrialpes.fr/index.php`.

[Web11c] WEBSITE: *Haptic Workstation*, May 2011. `http://www.cyberglovesystems.com/products/haptic-workstation/overview`.

[Web11d] WEBSITE: *Haption Virtuose*, May 2011. `http://www.haption.com/site/eng/html/materiel.php?item=1`.

[Web11e] WEBSITE: *icido GmbH*, May 2011. `http://www.icido.de`.

[Web11f] WEBSITE: *Memry Corporation*, May 2011. `http://memry.com/`.

[Web11g] WEBSITE: *MIT Kinect*, May 2011. `http://www.csail.mit.edu/videoarchive/research/hci/kinect-detection`.

[Web11h] WEBSITE: *Sensable Corp.*, May 2011. `http://www.sensable.com`.

[WLGP04] WAN, HUAGEN, YANG LUO, SHUMING GAO and QUNSHENG PENG: *Realistic virtual hand modeling with applications for virtual grasping.* In *Proceedings of the ACM SIGGRAPH International Conference on Virtual Reality Continuum* *and its Applications in Industry*, VRCAI, pages 81–87, New York, NY, 2004.

[WP09] WANG, ROBERT Y. and JOVAN POPOVIC: *Real-Time Hand-Tracking with a Color Glove*. ACM Transactions on Graphics, 28(3), 2009.

[WPFH98] WELLMAN, PARRIS S., WILLIAM J. PEINE, GREG FAVALORA and ROBERT D. HOWE: *Mechanical Design and Control of a High-Bandwidth Shape Memory Alloy Tactile Display*. In *Proceedings of the 5th International Symposium on Experimental Robotics*, pages 56–66, London, UK, 1998.

[WYHS02] WALAIRACHT, SOMSAK, KEITA YAMADA, SHOICHI HASEGAWA and MAKOTO SATO: *4 + 4 Fingers Manipulating Virtual Objects in Mixed-Reality Environment*. Presence, 11(2):134–143, 2002.

[WZ09] WELLER, RENE and GABRIEL ZACHMANN: *Inner Sphere Trees for Proximity and Penetration Queries*. In *Proceedings of Robotics: Science and Systems Conference*, RSS, Seattle (WA), USA, 2009.

[ZM98] ZHAI, SHUMIN and PAUL MILGRAM: *Quantifying Coordination in Multiple DOF Movement and Its Application to Evaluating 6 DOF Input Devices*. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Comuting Systems*, CHI, pages 320–327, 1998.

[ZR01] ZACHMANN, GABRIEL and A. RETTIG: *Natural and Robust Interaction in Virtual Assembly Simulation*. In *Proceedings of 8th ISPE International Conference on Concurrent Engineering: Research and Applications*, ISPE/CE, Anaheim (CA), USA, 2001.

Curriculum Vitae

Personal	Mathias Moehring
	Born on October 2nd, 1979 in Pasewalk, Germany

Education

2000 – 2005	Diploma Studies in Media Systems at Bauhaus-Universität Weimar, Germany
1992 – 1999	Secondary School “Max-Reinhardt-Gymnasium”, Berlin, Germany
1986 – 1992	Elementary School, Berlin, Germany

Professional

2011−now	IT-Consultant at Volkswagen AG Group IT, Wolfsburg Germany
2007 – 2011	Researcher at Volkswagen AG Group Research, Wolfsburg Germany
2005 – 2007	PhD Student at Volkswagen AG Group Research, Wolfsburg Germany

Wolfsburg, October 13, 2012

Publications

Moehring, Mathias and Froehlich, Bernd: *Pseudo-Physical Interaction with a Virtual Car Interior in Immersive Environments.* In *Proceedings of the 9th International Workshop on Immersive Projection Technology, 11th Eurographics Workshop on Virtual Environments*, IPT/EGVE, pages 181 − 189, Aalborg, Denmark, 2005.

Scheibe, Robert, Moehring, Mathias and Froehlich, Bernd: *Tactile Feedback at the Finger Tips for Improved Direct Interaction in Immersive Environments.* In *Proceedings of the IEEE Symposium on 3D User Interfaces*, 3DUI, pages 123 − 129, Charlotte (NC), USA, 2007.

Moehring, Mathias and Froehlich, Bernd: *Enabling functional validation of virtual cars through Natural Interaction metaphors.* In *Proceedings of IEEE Virtual Reality Conference*, VR, pages 27 − 34, Waltham (MA), USA, 2010 *(Best Student Paper Award).*

Moehring, Mathias and Froehlich, Bernd: *Effective Manipulation of Virtual Objects Within Arm's Reach.* In *Proceedings of IEEE Virtual Reality Conference*, VR, pages 131 − 138, Singapore, Singapore, 2011.

Moehring, Mathias and Froehlich, Bernd: *Natural Interaction Metaphors for Functional Validations of Virtual Car Models.* IEEE Transactions on Visualization and Computer Graphics (TVCG), 17(9) : 1195 − 1208, 2011.

Ehrenwörtliche Erklärung

Ich erkläre hiermit ehrenwörtlich, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Die aus anderen Quellen direkt oder indirekt übernommenen Daten und Konzepte sind unter Angabe der Quelle gekennzeichnet. Teile der Arbeit die bereits Gegenstand von Prüfungsarbeiten waren, sind ebenfalls unmissverständlich gekennzeichnet.

Weitere Personen waren an der inhaltlich-materiellen Erstellung der vorliegenden Arbeit nicht beteiligt. Insbesondere habe ich hierfür nicht die entgeltliche Hilfe von Vermittlung- bzw. Beratungsdiensten (Promotionsberater oder anderer Personen) in Anspruch genommen. Niemand hat von mir unmittelbar oder mittelbar geldwerte Leistungen für Arbeiten erhalten, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen.

Die Arbeit wurde bisher weder im In- noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt.

Ich versichere, dass ich nach bestem Wissen die reine Wahrheit gesagt und nichts verschwiegen habe.

Wolfsburg, October 10, 2012