# Effective Manipulation of Virtual Objects Within Arm's Reach

Mathias Moehring*
Group Research Virtual Technologies
Volkswagen AG

Bernd Froehlich†
Virtual Reality Systems Group
Bauhaus-Universität Weimar

## ABSTRACT

We present a study that compares finger-based direct interaction to controller-based ray interaction in a CAVE as well as in head-mounted displays. We focus on interaction tasks within reach of the users' arms and hands and explore various feedback methods including visual, pressure-based tactile and vibro-tactile feedback. Furthermore, we enhanced a precise finger tracking device with a direct pinch-detection mechanism to improve the robustness of grasp detection.

Our results indicate that finger-based interaction is generally preferred if the functionality and ergonomics of manually manipulated virtual artifacts has to be assessed. However, controller-based interaction is often faster and more robust. In projection-based environments finger-based interaction almost reaches the task completion times and the subjective robustness of controller-based interaction if the grasping heuristics relies on our direct pinch detection. It also improves significantly by adding tactile feedback, while visual feedback proves sufficient in head-mounted displays. Our findings provide a guideline for the design of fine grain finger-based interfaces.

**Index Terms:** B.4.2 [Input/Output Devices]: Channels and controllers H.3.4 [Systems and Software]: Performance evaluation H.5.2 [User Interfaces]: Interaction styles I.3.7 [3D Graphics and Realism]: Virtual Reality

## 1 INTRODUCTION

When developing virtual reality applications we often have the choice and the challenge to select the appropriate interaction methods. One interesting choice to make is the use of direct interaction by tracking body, hand and finger movements versus less direct input through controllers and buttons. 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. However, transferring the user's body into virtual environments can be challenging, particularly in projection-based virtual environments. On the other hand, less direct interaction through input devices and controllers can be very precise, robust and easy to learn, but it is limited with respect to the assessment of the functionality of manually operated mechanisms in virtual environments.

Pseudo-physical interaction metaphors provide a balanced compromise between robustness and realism for fine finger-based manipulation of constrained and unconstrained virtual objects [15]. While quite precise finger tracking is required and available, we found that the immediate and accurate recognition of finger pinches remains a challenge and limits the robustness of this approach. This is particularly the case for small and thin objects. In addition, there is the pertinent question of how to provide intuitive feedback once an object has been grasped if force feedback is not available.

The main goal of this work is to develop guidelines for selecting the appropriate input device and feedback methods for interaction tasks within the reach of the user's arms and hands. In our study we focus on the comparison of finger-based direct interaction to controller-based ray interaction in a CAVE as well as in head-mounted displays (HMD). We are interested in the relative performance and the positive and negative aspects of both approaches. To improve the robustness and feedback of finger-based interaction, we enhanced a precise optical finger tracking system with direct pinch-recognition capabilities as well as with tactile and vibro-tactile feedback.

Our extensive user study comparing direct and indirect interaction methods for a virtual car interior yields the following primary results:

- Finger-based interaction is clearly preferred over ray-based interaction for the assessment of various functionalities in a car interior.
- Ray-based interaction using a hand-held input device is faster than finger-based interaction. However, pinch-recognition allows finger-based interaction to be performed almost as quickly.
- Tactile feedback systems support the user with grasping and provide subjectively better interaction independent of the used display type.
- The pressure-based system is more comfortable, while the vibration-based system provides stronger feedback.
- Visual feedback alone can be sufficient for HMD-based applications.

The pinch-sensitive finger tracking device considerably enhances the grasping heuristics and provides robust and reliable grasping of all types of objects without impairing the judgment of the simulated real-world tasks caused by indirect interaction techniques. Our feedback methods provide a subjectively better interaction and considerably improve the capabilities of the users to judge if they have grasped an object. These results provide a guideline for the selection and implementation of realistic finger-based interactions in the automotive industry and many other areas.

## 2 RELATED WORK

Finding suitable metaphors and interfaces for interaction with virtual environments has been a major issue of VR-research. For our applications it is inevitable to use direct metaphors that provide realism and robustness. Common implementations of direct user interaction using the Virtual Hand metaphor [5] lack realism. They do not implement thorough grasping strategies based on fine-grain finger motion, but combine hand collision with a button press for selection and derive direct manipulation from the pose of the whole hand. We use pseudo-physical interaction as suggested in an earlier work [15], which has shown its potential for functional assessments of a car interior.

---

*e-mail: mathias.moehring@volkswagen.de
†e-mail:bernd.froehlich@uni-weimar.de

Metaphors based on physics simulations provide even more realistic object behavior including gravity and object-object collision. Borst and Indugula [3] realized such an interaction with a single object at a time, but until now physics engines have not been robust and all-encompassing enough to realize the broad range of object types and constraints found in our scenarios.

For common Virtual Hand implementations it is necessary, to provide a button press signal for selecting an object. The Pinch Glove uses cloth buttons to detect pinches between the user's fingers. For example, it was used for menu interaction, text input or navigation [6]. Van de Pol et al. attached tracking sensors to some of the fingers of a Pinch Glove and used it for direct picking, assuming an explicit pinch and the presence of the pinched fingers inside the bounding box of the desired object [22]. Cutler et al. realized two-handed interaction within arm's reach by using two tracked Pinch Gloves [9]. A combination of cloth buttons and a bend-sensing data glove was used by Laviola and Zeleznik [13] to enhance a variety of glove-based interaction techniques. In contrast, we want to combine optical finger tracking with pinch detection to enhance the robustness of our finger-based grasping heuristics – in particular for thin and tiny objects.

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 [4]. Other research examines and evaluates the particular characteristics of input devices [24, 23]. Poupyrev et al., amongst others, compared Virtual Hand and Virtual Pointer metaphors within the arm's reach, which is a scenario similar to ours. Their implementation of the direct metaphor does not provide realistic grasping. Instead, it resembles ray-based picking and thus it is not surprising that both techniques perform comparably well within the reach of the user's hands. McMahan et al. showed that controller-based interaction significantly outperforms natural interaction in the context of video games, while naturalism of interaction metaphors increases the perceived fun during the game [14]. For our 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.

Tactile feedback has been used as an information channel in a wide range of applications, such as tactile vests [11] and actuators that are mounted to input devices [17] or the user [18]. A comprehensive overview of tactile devices is provided by Benali-Khoudja et al. [2]. Usually vibro-tactile actuators are used because they are inexpensive, small and reliable [21]. The commercially available CyberTouch device has vibro-tactile actuators mounted on the back of the fingers of a common data glove. Unfortunately the location of the feedback is irritating since it is does not occur where the actual finger-object contact takes place. 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 [12]. Pabon et al. incorporated voice-coil motors for vibrational feedback into a data glove without performing any evaluation [16]. Various other devices provide tactile information at the fingers for sensory substitution [7].

An alternative to vibro-tactile stimuli is the use of pressure sensations for feedback. Scheibe et al. suggested the use of Shape Memory Alloy wires wrapped around the finger tips to provide tactile feedback during interaction [20]. Aoki et al. [1] have developed a similar feedback device but used a wire tightened by a motor for the feedback. However, 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 [20, 17, 18]. Our intention is to evaluate if a well-integrated tactile feedback system at the finger tips is capable of improving the objective and subjective interaction performance.

## 3 INPUT DEVICES AND FEEDBACK TYPES

We use an optical finger tracking system for the precise transfer of the user's hands into the VE, which is a prerequisite for direct finger-based interaction. This system can be enhanced to support the user during interaction tasks. With a mechanism capable of measuring pinches of the user's fingers, grasping reliability should be significantly improved. Tactile actuators added to the finger tracking can provide grasping information on the haptic channel and help the user by displaying the grasp status. The same support without using any additional hardware can be provided by visual feedback, wherein grasp information is demonstrated by material changes of the virtual hand. All the devices we used for the study were developed by the company A.R.T. GmbH and are explained in this section.

### 3.1 Finger Tracking System

The optical finger tracking system is the basis of all input devices evaluated in the study. It is an extension of the optical tracking system that is used for head tracking and is capable of a precise tracking of the user's hands and fingers. Therefore, the 6D pose of the hand and the 3D position of each finger are tracked by active markers (Figure 1). Plausible finger flections are estimated from the position of the finger tips in the hand tracking coordinate frame. Often we use a finger tracking device that only provides three fingers - thumb, index and middle finger – thereby reducing the dressing complexity of the system and remaining sufficient for direct interaction tasks. With this optical finger tracking system, the positions of the finger tips are tracked directly. Therefore, it provides the highest accuracy exactly where it is necessary for direct interaction – at the finger tips.

The tracking system reports the 6D pose of the hands, the position and orientation of each finger tip, the length of each finger segment and the angles between the segments. A 1D rotation is provided for both of the interphalangeal joints, while a 2D rotation describes the flection of the metacarpophalangeal joints. The palms of the hands are supposed to be rigid, considering that the determination of palm deformation is not possible with this system. Since the lengths of the finger segments are different for each user and play an important role for the inverse kinematics, they should be calibrated individually with a two-step calibration procedure before each application. The actual hand and finger poses are applied to a simple virtual hand model consisting of a rigid palm and fifteen finger segments per hand. Individual finger diameter and palm sizes cannot be calibrated properly, which is why common average sizes are used for all users.

### 3.2 Pinch-Sensitive Finger Tracking System

Some of the objects in the car interior are very thin or tiny. These objects are the most challenging objects for direct interaction metaphors because they are partly or even completely hidden by the user's hands. Furthermore, it is hard for some users to ensure proper collision of finger segments with these objects which is a prerequisite for valid grasps. Considering we use average finger diameters and due to the necessary finger thimble hardware it cannot be guaranteed that pinching fingers in reality leads to pinching virtual finger segments. On the other hand, 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. It is obvious that a grasp intention can be assumed when the user is pinching two or more fingers and at least one of the involved finger segments is colliding with a virtual object. With an extension of the finger tracking device these real finger pinches can be detected and then be used to enhance grasp detection.

For pinch detection, conductive stripes of metal are incorporated into the finger thimbles and the 6D hand target housing (Figure 1). By measuring currents it is possible to immediately detect finger

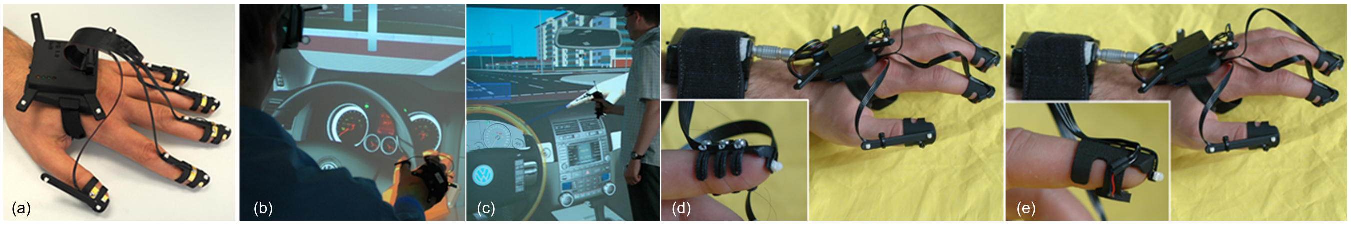


Figure 1: Pinch-sensitive device (a), visual feedback in the CAVE (b), indirect Flystick interaction in the CAVE (c), pressure-based (d) and vibration-based (e) tactile feedback devices with a close-up of the feedback actuators

pinches. Pinch states for each of the provided fingers can then be sent wirelessly to the VR application. Our grasping heuristics are enhanced with this pinch information. Whenever one of the involved finger segments is colliding with a virtual object and pinch state is positive this object is defined to be grasped by the finger segments that are involved in the pinch. Pinch information is only used for grasp detection, whereas decisions on grasp releases are based on the rules of the pseudo-physical interaction metaphor [15]. The pinch-sensitive system is used additionally to conventional grasp detection rules to improve grasp detection.

### 3.3 Visual Feedback

One way to provide feedback on grasping status is to color the grasping finger segments of the virtual hand representation to signal orange. This visual feedback gives the users the chance to explicitly see when 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 - totally immersing the user - visual feedback can be perceived easily by the users. However, since we use a simplified hand model with average finger and palm sizes, parts of the virtual hand can be seen in the CAVE-setup as well (Figure 1).

### 3.4 Pressure-Based Tactile Feedback System

Another way to provide grasping feedback is to use the haptic channel. Usually the haptic sense is not stimulated during our virtual reality applications. Although haptic rendering devices have been intensively researched, there are no devices available so far that are capable of providing sufficient haptic feedback for direct finger-based interaction applications. Therefore, the haptic channel is free to be used for displaying arbitrary tactile information. One possibility is to use tactile devices at the finger tips to give the user grasping feedback. Similar to the visual feedback implementation we provide feedback at those finger tips that are involved in valid grasps.

The pressure-based tactile feedback device (Figure 1) incorporates wires made from shape memory alloys (SMA) into the finger thimbles of the optical finger tracking system as it is described in [20]. Wires made from this material shorten when a low electrical current is applied and relax to their normal length afterwards. When wrapped around the finger tips they are able to apply a pressure sensation that can be clearly felt by the users. The wires are supplied and managed by a microcontroller incorporated into the housing of the 6D hand tracking target. It receives commands wirelessly by our VR application. The response time of the system is below $30ms$. The shortcoming of this kind of feedback is that the human sensoric system rapidly adapts to pressure sensations such that a continuous feedback cannot be displayed. On the other hand a pressure sensation provided at the finger tips perfectly matches the experience the users have from grasping in reality.

### 3.5 Vibration-Based Tactile Feedback System

A different kind of tactile feedback is realized with the vibration-based tactile feedback device (Figure 1). For this system, voice coil motors are incorporated into the finger thimbles instead of SMA wires. This vibration can be continuously felt by the users and thus 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 well when vibration sensations are used. Moreover, this kind of feedback is more obtrusive than the pressure sensation. The way the grasp feedback is implemented equals the pressure-based and visual feedback.

## 4 STUDY

With an extensive user study we evaluated the influence of input devices and feedback types on interaction task performance, subjective user judgment and user preference. We measured task completion times (TCT) complemented by subjective judgments of the users and a questionnaire quantifying the devices' influence on the direct interaction.

As a reference we implemented an indirect interaction metaphor for our interaction framework, which uses the Flystick (Figure 1), an input device commonly used for immersive industrial applications (e.g. as in assembly simulations). It is a tracked joystick-like handle with several buttons transmitting their state via radio. A virtual cursor – usually a 3D-model of an arrow with a ray attached to it – 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. 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.

For the study, we used both display types usually employed in our applications. On the one hand, projection-based systems like the CAVE [8] are used to virtually experience car interiors. Here the users are seated on very reduced mockups only providing a driver seat. The users always have their own body as a self reference. We use a three-sided high resolution CAVE providing 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 [15]. 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 [19] 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 comfort. The HMD used in our 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. We do not hinder the virtual hands 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.

## 4.1 Methods

### 4.1.1 Participants

Twelve subjects participated in our study, one being female the rest male. The ages ranged from 22 to 39 with two subjects being left-handed. All subjects had unimpaired or at least corrected sight. The group involved VR system experts as well as experienced VR users and undergraduate students considered as novice users.

### 4.1.2 Design

Our user study was conducted in a car interior evaluation scenario typical for our applications. As a test case, we defined five interaction tasks covering a broad range of typical object properties, including the lowering of the driver's sun shield, the opening of the driver's door, the adjustment of the interior mirror, putting a freely movable soda bottle from passenger's footwell to the center console and the turning of the light switch.

All objects except the soda bottle have constraints: the sun shield, the door and the light switch have one rotational degree of freedom and the interior mirror can be freely rotated around a fixed center. Due to their location two of the objects are manipulated with the left hand (e.g. door, light switch). There are thin and small objects (e.g. sun shield, light switch) as well as a bigger object (e.g. soda bottle). The objects have to be grasped explicitly; just pushing objects around is disabled.

The procedure of each interaction task consists of object selection (grasping), manipulation from design position to a pre-defined target location and deselection (release). The target position is visualized as a semi-transparent copy of the object (Figure 2). We attached reference points to the interactive objects to facilitate the task completion decision. The interaction task is completed if the distance between the reference points attached to the interactive object and those attached to the target visualization is below a given tolerance. For task completion it is not necessary that this condition remains fulfilled after object release.

We included visual feedback for task completion. Therefore, a green bounding sphere appears for a short time around the car interior. For start and stop of the interaction task we included a start/stop region at the center of the steering wheel. It is visualized as a semi-transparent sphere (r=80mm) which becomes opaque if the user's index finger tip - according to the handedness of interaction task - is inside the tolerance region. TCT measurement starts when the tolerance region is left by the user and stops when it is re-entered. The interaction scenario can be seen in Figure 2.

Figure 2: Interaction scenario with the user performing the sun shield task

Table 1: Conditions of the study

| | | CAVE | HMD |
|---|---|---|---|
| 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 |

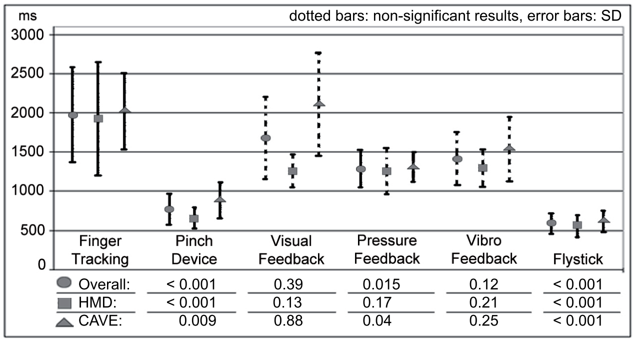
### 4.1.3 Procedure

For our evaluation we combined the measurement of objective task completion times with subjective measures and a questionnaire concerning several interaction aspects and device properties. We had twelve conditions that were evaluated, combining two display systems with six input devices (Table 1). Each of the twelve users performed the interaction tasks in each of the twelve conditions. For simplification of the study process, we divided the participants into two groups. One started evaluating the six input devices in the CAVE, while the other 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. To simplify the study flow and reduce the number of calibrations and device changes we used the five-finger pinch-sensitive device for all FT-, PSD- and VF-conditions. For FT- and VF-conditions we switched off the pinch analysis and for VF-conditions we added visual feedback in our scene. Please note, that the tactile feedback systems do not provide thimbles for the ring and pinky finger, which play a minor role during grasping interaction.

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 we did not want to evaluate differences among the objects. We encouraged the users to familiarize themselves with each condition beforehand by interacting freely with all objects. After each interactional task - conducted three times in sequence - we asked the users to judge this particular interaction on a seven item Likert scale with one being "very poor" and seven meaning "very good". Finger-based interaction tasks are influenced by many factors that do not necessarily depend on the input devices, such as tracking reliability and individual user performance. After all interaction tasks of one condition the participants filled out a questionnaire with questions concerning individual device and interaction properties (Table 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, we

Table 2: Device and interaction properties (Q1-10) and user preferences (Q11-14) to be rated in the questionnaire

| ID: Question | ID: Question |
|---|---|
| Q1: Device unobtrusiveness | Q8: Feedback unobtrusiveness |
| Q2: Intensity of the feedback | Q9: Interaction realism |
| Q3: Comfort of the feedback | Q10: Car judgment capabilities |
| Q4: Grasping ability | Q11: Preference to use the PSD |
| Q5: Ability to recognize a grasp | Q12: Pressure vs. vibro feedback |
| Q6: Object releasing ability | Q13: Visual vs. tactile feedback |
| Q7: Immediateness of feedback | Q14: Feedback vs. no feedback |



| | | | | | |
|---|---|---|---|---|---|
| ● Overall: | < 0.001 | 0.39 | 0.015 | 0.12 | < 0.001 |
| ■ HMD: | < 0.001 | 0.13 | 0.17 | 0.21 | < 0.001 |
| ▲ CAVE: | 0.009 | 0.88 | 0.04 | 0.25 | < 0.001 |

Figure 3: Mean intra-subject standard deviations for all conditions. The p-values below each condition refer to the statistical significance of the difference to the bare finger tracking conditions.

completed the study with preference questions. Here the participants had to directly choose one of two options (Table 2).

If an interaction task is too difficult for the user in one condition, this particular task is 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 is noted. In most of the cases significance levels were calculated by a paired t-test. If data sets were not normally distributed (Shapiro-Wilk test) we performed a Wilcoxon rank-sum test instead. The global significance level of $p = 0.05$ was adjusted with the Holm-Bonferroni method [10] for multiple tests.

## 4.2 Results and Discussion

Through our study, we want to evaluate four main issues. First, we want to know how the direct interaction metaphor with the bare finger tracking device (without any enhancements) performs compared to the reference input device, the Flystick. Second, we want to quantify if the pinch-sensitive finger tracking device enhances interaction capabilities. Third, since the visual feedback does not need any complementary hardware it would be an easy way to help the user with interaction if this support could be shown. Finally, we wanted to know if the tactile feedback devices make interaction easier and more robust. We structure this section according to these main questions.

### 4.2.1 Direct vs. Indirect Interaction

With our study we first compared the task performance and subjective judgment of direct interaction using bare finger tracking with our reference metaphor implementing indirect interaction with the Flystick. On the one hand we expect the Flystick method to provide easier and more robust interaction. On the other hand we expect this abstract metaphor to provide reduced interaction realism and therefore less support for functional validations of the virtual car.

Our study clearly confirmed our expectations. The indirect metaphor performed considerably better than unsupported direct interaction. Task completion times significantly differed and were considerably faster for all Flystick interactions. A significant difference was also observed in the FL-C as well as in the FL-H condition (Figure 4). This quantitative benefit is also reflected by subjective interaction ratings (Figure 4). Furthermore, the interaction with the Flystick was more robust leading to significantly lower intra-subject variance of task performance (Figure 3).

As we expected, the advantage of the indirect interaction is that it makes the grasping very easy — only pointing and pressing a button is required. We could see that the very fast and subjectively better Flystick interaction correlated with a significantly higher subjective rating for all grasping questions (Figure 5). All these effects could be found equally in both display systems, while no display-related effects could be encountered.

Although higher task performance and user ratings can be achieved with indirect interaction, we believe that a realistic interaction needs direct input. This also was confirmed by the questionnaire. The participants rated controller-based interaction considerably less realistic. While the Flystick interaction has to be considered to be an abstraction, direct interaction provides a comparatively more realistic interaction (Figure 5). With higher levels of realism, the properties of a virtual car interior can be better evaluated. Consequently, subjective ratings for the question on the ability to judge the virtual car are higher for direct interaction (Figure 5). Interestingly the difference in the CAVE was not as considerable as it was in the HMD and thus was not significant. We expected this display-related effect. 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 techniques.

### 4.2.2 Pinch-Sensitive Finger Tracking 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 and we use this information to improve grasp detection. The users were not informed about the special character of this input device before the trials. In this section, we compare the pinch-sensitive device with bare finger tracking. Especially for the challenging objects we expect the interaction to be improved comparably to controller-based interaction. We expect this improvement to be achieved without any loss of realism since the same direct interaction technique is used as it is with the bare finger tracking. Actually it has to be expected that - compared to the bare finger tracking - higher interaction performance and higher subjective ratings consequently lead to higher realism ratings.

As we expected, a clear interactional improvement can be obtained with the pinch-sensitive device. The achieved task completion times are only 10 to 20 percent slower than those we have seen with indirect interaction, but they are significantly better than those of the bare finger tracking. Again the higher task performance is reflected by the subjective ratings given by the subjects after each interaction. As with indirect interaction a higher robustness was indicated by lower intra-subject standard deviations of task performance compared to the bare finger tracking. All these effects could be found equally in both display systems without any significant effect related to the displays (Figure 4 and 3).

Interestingly the benefit for task performance of the pinch-sensitive system could not be equally found for all interaction tasks. 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.

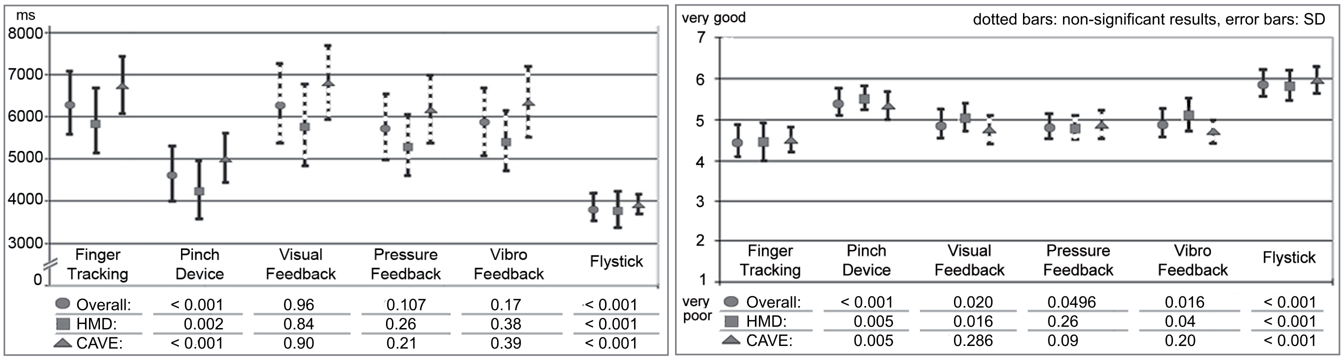


Figure 4: Mean task completion times and subjective user ratings for all conditions. The p-values below each condition refer to the statistical significance of the difference to the bare finger tracking conditions.

Thus, pinch detection did not improve grasp performance (interior mirror: $p = 0.53$, bottle: $p = 0.21$). Grasping for the other three objects was facilitated by pinch detection since the subjects almost always employed a pinch grasp (sun shield: $p < 0.001$, door: $p < 0.001$, light switch: $p < 0.001$). Especially the light switch which was, due to its minor size a problem for most of the users, could be grasped considerably better with pinch detection.

Obviously the pinch-sensitive device helps the users during direct interaction by significantly improving grasping. The users were able to grasp objects considerably better (Figure 5). We were surprised to see that the users even felt significantly supported with grasping judgment and object release, only releasing in the CAVE did not show significant improvement (Figure 5). Both issues are not explicitly supported by the device, especially releasing follows the same rules as they are used for the bare finger tracking. We believe that the overall better interaction with the pinch-sensitive device leads to this inexplicable improvement.

For our 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 normal finger tracking device. In fact, all PSD-conditions are rated considerably higher than the FL-conditions for Q9 and Q10 (Figure 6). Taking the whole sample into account, the user ratings on realism under PSD-condition were even significantly better than in FT-conditions (Figure 5). In contrast to the bare finger tracking, better car judgment capabilities compared to the indirect interaction were found in both display systems with a weak effect in the CAVE. 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 its ease of use are important for realistic interaction metaphors.

### 4.2.3 Visual Feedback

Another way to support direct interaction is to provide grasping feedback. Conventionally only the beginning motion of a grasped object is an indication for grasp evolution. For the visual feedback we changed the color of the virtual hands' finger segments defining a valid grasp. The disadvantage of this approach is that the virtual hand mostly disappears behind the real hand in projection-based virtual environments. We expect 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, we expect significant differences when CAVE- and HMD-conditions are compared. 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.

Our study results reveal that visual feedback is only supporting the user subjectively. And as we expected, this improvement can only be proven for the HMD-condition (Figure 4). Further explanations for the differences related to the display are provided by the results of the questionnaire. The visual feedback is perceived stronger in the HMD as it is in the CAVE (Q2) and the support with grasp judgment is significantly higher (Figure 6). Compared to the bare finger tracking, the visual feedback provides this intended improvement for the whole sample and in the HMD-condition; for the CAVE-condition no significant effect was separately encountered (Figure 5). A weak but significant support of the feedback can be also found for the actual grasping, but again this effect was not proven in the projection-based environment (Figure 5). Here the difference between HMD- and CAVE-condition was not significant. As we expected, these results verify that a visual feedback at the virtual hand is only sufficient if a Head Mounted Display is used. The benefit of the feedback for actual grasping support (Q4) is lower than for grasping judgment (Q5). This is due to the fact that we provided grasping feedback informing the user when a valid grasp evolved instead of giving collision feedback that eventually would help the user to actually grasp an object.

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 hardware (Q1) higher than for the PTF conditions in the CAVE (with a weak significant effect) and if all trials are considered (Figure 6). The higher ratings compared to the vibration-based feedback did not yield a significant effect. In the next section, we further evaluate the difference of PTF and VTF concerning the comfort ratings. 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 (Q8) 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 (Figure 6).

### 4.2.4 Tactile Feedback

Tactile feedback systems use the haptic channel to provide grasping information. The pressure-based device uses thin 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, we expect these systems to support the user with grasping judgment and therefore to provide a better interaction. Since these devices use the haptic channel, we do not expect significant differences depending on the display used. The devices are more obtrusive compared to the bare finger tracking, which should be reflected by lower hard-

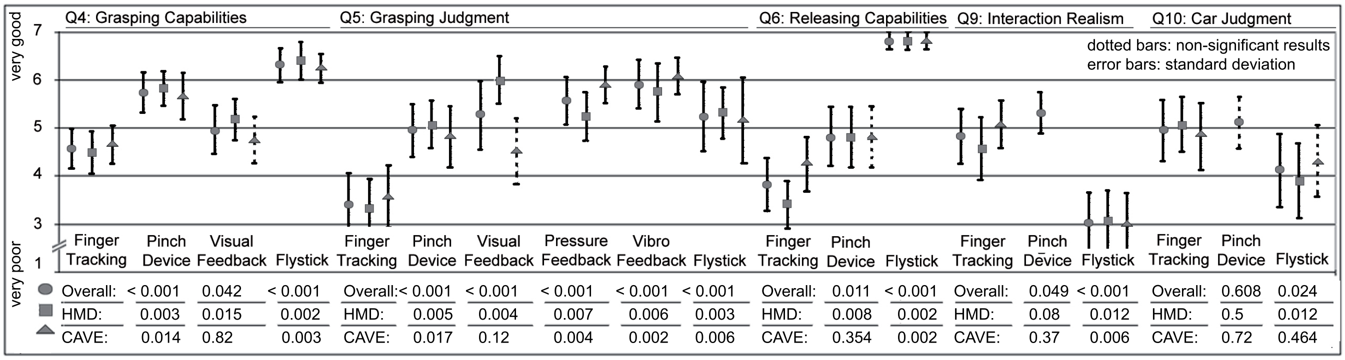


Figure 5: Selected results from the questionnaire. Mean subjective ratings for several questions and conditions. The p-values below each condition refer to the statistical significance of the difference to the respective results of the bare finger tracking conditions.

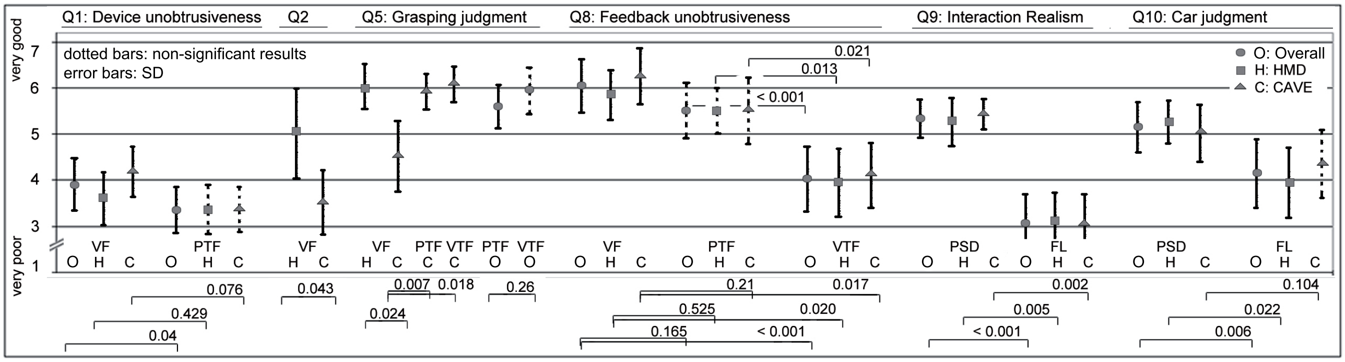


Figure 6: Further results of the questionnaire. Mean subjective ratings for several questions and conditions are directly compared. The p-values refer to the statistical significance of the differences.

ware comfort ratings. And finally, we expect the pressure-based system to provide less disturbing feedback than the vibration-based system.

Similar to the visual feedback task performance is hardly improved by tactile feedback; only a tendency can be encountered for the pressure-based system. Subjectively, both systems support the users during interaction and – indicated by intra-subject task performance standard deviations – robustness is improved by them as well, significantly with the pressure-based feedback (Figure 4 and 3). In general the tactile feedback systems do not improve interaction as much as the pinch-sensitive system.

Looking at the results of the questionnaire, a strong support of the ability to judge grasps can be seen for both systems in each display system (Figure 5). 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 (Figure 6). The subjective rating of the actual grasping is hardly improved by tactile grasping feedback. Again this is due to the fact that we did provide grasping feedback instead of collision feedback.

Concerning the comfort ratings, differences between the feedback devices can be identified. We would have expected that the pressure-based feedback earns lower hardware obtrusiveness ratings (Q1) because of the openly mounted wires. This prejudice was not confirmed by our study, considering no differences can be found when Q1-ratings of PTF and VTF are compared. In contrast, we expected the pressure-based system 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 (Figure 6). In summary, both tactile feedback systems support the user with grasping and provide subjectively better interaction independent of the used display type. While the vibration-based system provides stronger feedback, the pressure-based system is more comfortable.

## 4.3 User Preference and Summary

Our study proved a number of effects that we expected from the individual characteristics of each input and output device. Generally, it has to be stated that an indirect interaction technique in combination with the clear feedback of a button press provided by a typical hand-held input device provides the highest performance and most robust interaction with virtual objects. On the other hand, this kind of interaction lacks of realism, thus making it hard to evaluate this virtual interaction and transfer results obtained with it into reality. While users are able to compensate this disadvantage in the CAVE scenario, this is very hard in the HMD. A realistic interaction is only possible with direct techniques.

Since the realistic interaction does require direct input with finger tracking, the question is if this direct interaction can be enhanced by further improvements to the finger tracking device or by any kind of additional feedback. With our study, we have shown that the pinch detection mechanism attached to the bare finger tracking device clearly improved interaction with the virtual car. In contrast to the usage of indirect input, this grasping improvement is not paired with reduced realism or car judgment abilities. On the contrary, the benefit concerning robustness and ease-of-use reflects in higher realism ratings compared to the bare finger tracking. Consequently, eleven out of twelve users preferred to use the pinch-sensitive system (Q11) (with PSD: 91.67%, without: 8.33%).

Concerning feedback, we found out that it supports interaction at least in a subjective way. The interaction benefit obtained with the grasping feedback was not as clear as it was for the pinch-sensitive device or the indirect interaction. It would be interesting if collision feedback would provide more interactional improvement than the grasping feedback we realized. With this kind of feedback, the user might be supported with the actual grasping instead of grasping judgment. It is unclear which kind of feedback provides better interactional support. However, ten out of twelve users preferred to be supported by any kind of feedback (Q14) (with feedback 83.33%, without feedback 16.67%). For HMD-applications, obviously it is sufficient to use visual grasping feedback at the virtual hand that is less obtrusive for the users. Consequently, user preference is quite balanced concerning the kind of feedback (Q13) (visual: 41.67%, tactile: 58.33%). When projection-based immersive environments are used, this kind of feedback is useless and tactile feedback has to be used. Since the user preferences were questioned after all trials, the users probably decided with having the visual feedback in the HMD in mind where it worked best. Comparing the tactile devices, the vibration-based system seems to provide more feedback paired with a lack of comfort. This leads to a tendency for preferring the pressure-based system (Q12) (PTF: 66.67%, VTF: 33.33%).

## 5 CONCLUSIONS AND FUTURE WORK

The results of our extensive user study clearly show that finger-based interaction is preferred over indirect interaction for the assessments of various functionalities in 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 car interior. The deficits of direct methods concerning performance and robustness can be almost compensated by combining pinch-recognition with a precise finger tracking system. Grasping feedback is a requirement to judge grasp status. It is not sufficient to just have an object follow the user's hand motion once it is grasped. While visual feedback alone is mostly 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. These observations enable the robust and reliable implementation of interaction metaphors for realistic finger-based interactions.

A combination of pinch detection and tactile feedback has the potential to further improve direct interaction and increase its acceptance. A hybrid device provides robust grasping by accurate pinch detection and clearly communicated interaction states by the feedback. However, it remains challenging to integrate both mechanisms into the thimbles of the finger tracking device. While we use the feedback to report the grasp status to the users, an alternative would be to communicate finger collisions to help the user localize the actual grasp location. Further evaluation on the challenges and shortcomings of both approaches remains necessary. The vibration-based tactile feedback could be further improved to provide tactile patterns representing different states throughout the grasping process, which could also be achieved by a combination of different feedback methods. Unless unobtrusive, precise and crisp full finger haptic feedback becomes available, our finger-based interfaces provide a good and robust compromise for the realistic assessment of various functionalities of newly designed objects in virtual environments.

## REFERENCES

[1] T. Aoki, H. Mitake, D. Keoki, S. Hasegawa, and M. Sato. Wearable haptic device to present contact sensation based on cutaneous sensation using thin wire. In *Proceedings of ACM ACE'09*, Athens, Greece, 2009.

[2] M. Benali-Khoudja, M. Hafez, J. M. Alex, and A. Kheddar. Tactile interfaces: a state-of-the-art survey. In *Proceedings of International Symposium on Robotics*, pages 721–726, 2004.

[3] C. W. Borst and A. P. Indugula. A spring model for whole-hand virtual grasping. *Presence*, 15(1):47–61, 2006.

[4] D. A. Bowman and L. F. Hodges. An evaluation of techniques for grabbing and manipulating remote objects in immersive virtual environments. In *Proceedings of Symposium on Interactive 3D Graphics (SI3D'97)*, pages 35–38, 182, 1997.

[5] D. A. Bowman, E. Kruijff, J. J. LaViola, and I. Poupyrev. *3D User Interfaces: Theory and Practice*. Addison Wesley Longman Publishing Co., Inc., Redwood City, CA, USA, 2004.

[6] D. A. Bowman, C. A. Wingrave, J. M. Campbell, V. Q. Ly, and C. J. Rhoton. Novel uses of pinch gloves for virtual environment interaction techniques. *Virtual Reality*, 6:122–129, 2002.

[7] G. C. Burdea. *Force and Touch Feedback for Virtual Reality*. John Wiley & Sons, Inc., New York, NY, USA, 1996.

[8] C. Cruz-Neira, D. J. Sandin, and T. A. DeFanti. Surround-screen projection-based virtual reality: The design and implementation of the cave. In *Proceedings of ACM SIGGRAPH'93*, pages 135–142, 1993.

[9] L. D. Cutler, B. Fröhlich, and P. Hanrahan. Two-handed direct manipulation on the responsive workbench. In *Proceedings of Symposium on Interactive 3D Graphics (SI3D'97)*, pages 107–114, 1997.

[10] S. Holm. A simple sequentially rejective multiple test procedure. In *Scandinavian Journal of Statistics*, pages 65–70, 1979.

[11] L. A. Jones, M. Nakamura, and B. Lockyer. Development of a tactile vest. In *Proc. of IEEE Haptics'04*, pages 82–89, Chicago, IL, 2004.

[12] P. Kammermeier, A. Kron, J. Hoogen, and G. Schmidt. Display of holistic haptic sensations by combined tactile and kinesthetic feedback. *Presence*, 13(1):1–15, 2004.

[13] J. Laviola and R. C. Zeleznik. Flex and pinch: A case study of whole hand input design for virtual environment interaction. In *Proc. of Intl. Conf. on Computer Graphics and Imaging*, pages 157–168, 1999.

[14] R. McMahan, A. Alon, S. Lazem, R. Beaton, D. Machaj, M. Schaefer, M. Silva, A. Leal, R. Hagan, and D. Bowman. Evaluating natural interaction techniques in video games. In *Proceedings of IEEE Symposium on 3D User Interfaces (3DUI'10)*, pages 11–14, Waltham, MA, 2010.

[15] M. Moehring and B. Froehlich. Enabling functional validation of virtual cars through natural interaction metaphors. In *Proceedings of IEEE Virtual Reality Conference (VR'10)*, pages 27–34, Waltham, MA, 2010.

[16] S. Pabon, E. Sotgiu, R. Leonardi, C. Brancolini, O. Portillo-Rodriguez, A. Frisoli, and M. Bergamasco. A data-glove with vibro-tactile stimulators for virtual social interaction and rehabilitation. In *10th Annual Intl Workshop on Presence*, Barcelona, Spain, 2007.

[17] H. Regenbrecht, J. Hauber, R. Schönfelder, and A. Maegerlein. Virtual reality aided assembly with directional vibro-tactile feedback. In *Proceedings of ACM GRAPHITE'05*, pages 381–387, Dunedin, New Zealand, 2005.

[18] J. Ryu and G. J. Kim. Using a vibro-tactile display for enhanced collision perception and presence. In *Proceedings of ACM VRST'04*, pages 89–96, Hong Kong, China, 2004.

[19] H. Salzmann and B. Froehlich. The two-user seating buck: Enabling face-to-face discussions of novel car interface concepts. In *Proceedings of IEEE Virtual Reality Conference (VR'08)*, pages 75–82, Reno, NV, 2008.

[20] R. Scheibe, M. Moehring, and B. Froehlich. Tactile feedback at the finger tips for improved direct interaction in immersive environments. In *Proceedings of IEEE Virtual Reality Conference (VR'07)*, pages 293–294, Charlotte, NC, 2007.

[21] K. B. Shimoga. A survey of perceptual feedback issues in dexterous telemanipulation: Part ii. finger touch feedback. In *Proceedings of IEEE Virtual Reality Conference (VR'93)*, pages 271–279, 1993.

[22] R. van de Pol, W. Ribarsky, L. Hodges, and F. Post. Interaction in semi-immersive large display environments. In *Proceedings of EGVE'99*, pages 157–168, 1999.

[23] C. Ware. Using hand position for virtual object placement. *The Visual Computer*, 6(5):245–253, 1990.

[24] S. Zhai and P. Milgram. Quantifying coordination in multiple dof movement and its application to evaluating 6 dof input devices. In *Proceedings of ACM CHI'98*, pages 320–327, 1998.