# Audio-Tactile Proximity Feedback for Enhancing 3D Manipulation

Alexander Marquardt, Ernst Kruijff, Christina Trepkowski, Jens Maiero, Andrea Schwandt, André Hinkenjann
Institute of Visual Computing,
Bonn-Rhein-Sieg University of Applied Sciences
Sankt Augustin, Germany
alexander.marquardt@h-brs.de

Wolfgang Stuerzlinger
School of Interactive Arts and Technology, Simon Fraser University
Surrey, Canada
w.s@sfu.ca

Johannes Schoening
University of Bremen
Bremen, Germany
schoening@uni-bremen.de

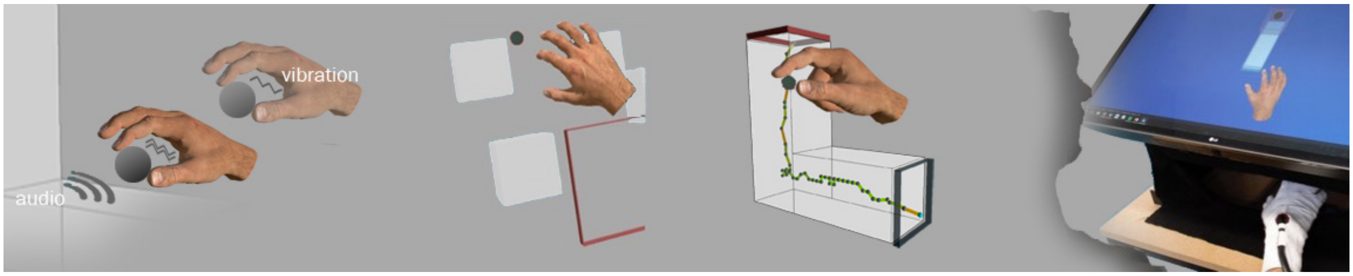

**Figure 1: From Left to right: Schematic representation of proximity-based feedback, where directional audio and tactile feedback increases in strength with decreasing distance, scene exploration task study 1, tunnel task study 2 with example path visualization (objects in study 1 and 2 were not visible to participants during the experiments), and reach-in display with the tunnel (shown for illustration purposes only).**

## ABSTRACT

In presence of conflicting or ambiguous visual cues in complex scenes, performing 3D selection and manipulation tasks can be challenging. To improve motor planning and coordination, we explore audio-tactile cues to inform the user about the presence of objects in hand proximity, e.g., to avoid unwanted object penetrations. We do so through a novel glove-based tactile interface, enhanced by audio cues. Through two user studies, we illustrate that proximity guidance cues improve spatial awareness, hand motions, and collision avoidance behaviors, and show how proximity cues in combination with collision and friction cues can significantly improve performance.

## KEYWORDS

Tactile feedback; 3D user interface; hand guidance

**ACM Reference Format:**
Alexander Marquardt, Ernst Kruijff, Christina Trepkowski, Jens Maiero, Andrea Schwandt, André Hinkenjann, Wolfgang Stuerzlinger, and Johannes Schoening. 1997. Audio-Tactile Proximity Feedback for Enhancing 3D Manipulation. In *Proceedings of (VRST'18)*. ACM, New York, NY, USA, Article 4, 10 pages. https://doi.org/10.475/123_4

## 1 INTRODUCTION

Despite advances in the field of 3D user interfaces, many challenges remain unsolved [32]. For example, it is still difficult to provide high-fidelity, multisensory feedback [30]. However, as in real-life, there are many tasks that depend on multisensory cues. For example, in complex or dense scenes, 3D interaction can be difficult: hand motions are hard to plan and control in the presence of ambiguous or conflicting visual cues, which can lead to depth interpretation issues in current unimodal 3D user interfaces. This, in turn, can limit task performance [32]. Here, we focus on 3D manipulation tasks in complex scenes. Consider a virtual reality training assembly procedure [6], in which a tool is selected and moved through a confined space by hand, and then using the tool to turn a screw. Here, multiple visual and somatosensory (haptic) cues need to be integrated to perform the task. A typical problem during manipulation in unimodal interfaces in such scenarios is hand-object penetration, where the hand passes unintendedly through an object. Such object penetrations can occur frequently, especially when users cannot accurately judge the spatial configuration of the scene around the hand, making movement planning and correction difficult. However, similar to real-world scenarios, multisensory cues can disambiguate conflicting visual cues, optimizing 3D interaction

---

*VRST'18, Nov.28-Dec.01, Tokyo, Japan*

ACM ISBN 123-4567-24-567/08/06.
https://doi.org/10.475/123_4

performance [53]. Cues can be used proactively and adaptively, affording flexible behaviour during task performance [53].

## 1.1 Motor Planning and Coordination

Planning and coordination of selection and manipulation tasks is generally performed along a task chain with key control points. These control points typically relate to contact-driven biomechanical actions [22]. As such, they contain **touch** cues that relate to events about touching objects to select them (selection) or move along a trajectory (manipulation). This may contain various hand **motion** and **pose** actions that are performed within the scene context, e.g., for steering the hand during manipulation tasks. There should be sufficient indication as to where the hands touches objects upon impact (collision contact points) or slides along them (friction contact points), while other indications, such as object shape or texture, can also be beneficial [24].

Multisensory stimuli enable learning of sensorimotor correlations that guide future actions, e.g., via corrective action patterns to avoid touching (or penetrating) an object [22]. In real-life, to steer hand motions and poses, we depend typically on visual and physical constraints. E.g., lightly touching a surrounding object might trigger a corrective motion. However, manipulation tasks are also performed independent of touch cues, namely through self-generated proprioceptive cues [38]. Such cues may have been acquired through motor learning [47]. Although not the main focus of this work, motor learning can be an important aspect for skill transfer between a 3D training application and the real-world [11, 28], thereby potentially also "internalizing" proprioception-driven actions for later recall.

## 1.2 Research questions

Our novel guidance approach, which is described in more detail in section 3, is based on audio-tactile proximity feedback to communicate the direction and distance of objects surrounding the user's hand. Feedback is used to plan and coordinate hand motion in 3D scenes. Our research is driven by the following research questions (RQs) that assess how we can guide the hand **motion** before and during 3D manipulation tasks using such feedback.

*RQ1. Do scene-driven proximity cues improve spatial awareness while exploring the scene?*

*RQ2. Can hand-driven proximity cues avoid unwanted object penetration or even* ***touching*** *proximate objects during manipulation tasks?*

In this paper, we measure the effect of proximity cues in combination with other haptic cue types (in particular collision and friction). Towards this goal, study 1 (scene exploration) explores the general usefulness of proximity cues for spatial awareness and briefly looks at selection, while study 2 looks specifically at the effect of proximity on 3D manipulation tasks. In our studies, we specifically look at touch and motion aspects, while leaving support for pose optimization as future work. As a first step, we focus on feedback independently of visual cues, to avoid confounds or constraints imposed by such cues.

## 1.3 Contributions

Our research extends previous work by Ariza et al. [3] that looked into low resolution and non-directional proximity feedback for 3D selection purposes. We provide new insights into this area of research by looking at higher-resolution and directional cues for manipulation (instead of selection) tasks. Our studies illustrate the following benefits of our introduced system:

- In the scene exploration task, we show that providing proximity feedback aids spatial awareness through a higher number of tactors (18 vs. 6), which improves both proximity feedback (20.6%) and contact point perception (30.6%). While the latter is not unexpected, the results indicate the usefulness of a higher-resolution tactile feedback device.
- We illustrate how the addition of either audio or tactile proximity cues can reduce the number of object collisions up to 30.3% and errors (object pass-throughs) up to 56.4%.
- Finally, while friction cues do not show a significant effect on measured performance, subjective performance ratings increase substantially, as users thought that with friction (touch) they could perform faster (18.8%), more precisely (21.4%), and react quicker to adjust hand motion (20.7%).

# 2 RELATED WORK

In this section, we outline the main areas of related work. **Haptic feedback** has been explored for long, though is still limited by the need for good cue integration and control [30, 50], cross-modal effects [41], and limitations in actuation range [18]. The majority of force feedback devices are grounded (tethered). Such devices are often placed on a table and generally make use of an actuated pen that is grasped by the finger tips, e.g., [54]. Only few glove or exoskeleton interfaces exist that enable natural movement, while still providing haptic feedback, such as grasping forces, e.g., [7]. In contrast, tactile methods remove the physical restrictions of the aforementioned actuation mechanisms, and thus afford more flexibility, by substituting force-information in tactile cues, not only for 3D selection and manipulation tasks [25, 31], but also for other tasks like navigation [29]. In 3D applications, recent research looked at smaller, handheld (e.g. [5]) or glove-based (e.g. [15, 48]) tactile actuators [2, 9]. Instead of stimulating only the finger tips and inner palm using a limited number of tactors, researchers have also looked into higher-density grids of vibrotactors to stimulate different regions of the hand [16, 36, 45], but these approaches are currently limited to localized areas.

Some researchers have explored **proximity feedback** with a haptic mouse [19], using vests for directional cues [33], to trigger actions [4], and for collision avoidance using audio feedback [1]. Most relevant to our tactile proximity feedback is a system called SpiderSense [37], which uses tactors distributed over the body to support navigation for the visually impaired. This kind of feedback is similar to a distance-to-obstacle feedback approach [17] and a glove-based approach for wheelchair operation [52]. Furthermore, tactile guidance towards a specific target [40] or motion and pose [35] has shown promise. Yet, both the usage context and approaches differ fundamentally from our tactile guidance approach, which aims to increase spatial awareness to better support manipulation of objects in 3D interaction scenarios. Finally, Ariza et al. studied non-directional feedback for selection tasks, showing that different types of feedback affect the ballistic and correction phases of selection movements, and significantly influence user performance [3].

## 3 CHALLENGES

Providing multisensory cues – in particular haptics – to complement visual-only feedback has benefits for 3D manipulation tasks. However, while haptic cues aid in guiding hand motion and poses, their inclusion in 3D user interfaces is challenging. Traditional grounded haptic interfaces (force feedback devices) provide cues that support the user in performing fine manipulation tasks, for example by guiding the hand by constraining its' motion. As such, haptics potentially ameliorate any negative effect of visual ambiguities [8] and has been shown to improve selection tasks [10]. However, haptic devices often have limitations, such as operation range, the kind of contact information being provided, and issues related to the type of the used feedback metaphor. For example, popular actuated pen devices, such as the (Geomagic) Phantom, do not necessarily comply to how users perform actions in the real world, as they support only a pen grip instead of full-hand interaction. Such interfaces do not provide contact point feedback across the full hand, which limits the feedback that users can use to plan and coordinate selection and manipulation tasks: users will be unaware where the hand touches another object, even though this information may be required to steer hand motion and poses. While full-hand interfaces exist, they are often expensive, have mostly a limited operation range, and can be cumbersome to use.

Tactile interfaces are an interesting alternative to traditional grounded haptic (force feedback) devices, as they provide portable solutions with good resolution and operation range [55]. However, designing effective tactile cues is challenging, as haptic (force) stimuli cannot be fully replaced by tactile cues without loss of sensory information [25]. Furthermore, simulating contact has its limitations, as untethered systems cannot restrict physical motion. As a result – similar to visual-only feedback conditions – users may still pass through virtual objects unintentionally, as users often cannot react quickly (enough) to tactile collision cues [12]. During selection tasks, and before colliding (selection) with an object, the hand typically goes through a fast (ballistic) motion phase, followed by fine, corrective motor actions [34]. Similarly, once the hand touches an object in the scene during a manipulation task, a corrective movement may be performed, e.g., to steer away from this object. However, as movement is typically not perfect, the users' hand will often move into or through the object even though a tactile collision cue is perceived, especially when a corrective movement is initiated too late. The presence of any depth perception issues or other visual ambiguities typically make this situation only worse. During selection tasks, this may for example lead to overshooting [3]. Furthermore, especially for thin objects, users may move (repeatedly) through the object during manipulation, as such objects trigger only short bursts of collision feedback.

## 4 APPROACH

We aim to overcome limitations associated with the untethered nature of many tactile devices – in particular the inability to constrain human motion – by guiding the hand through proximity feedback.

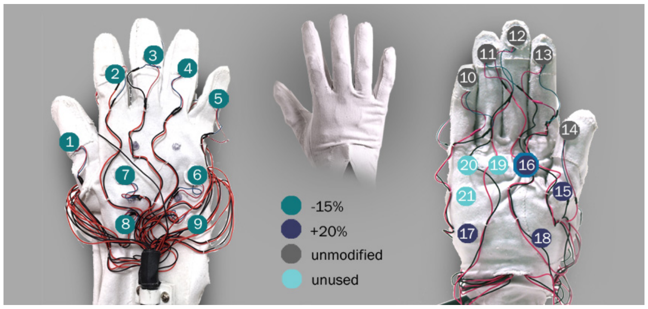

**Figure 2: Tactor IDs and balancing of our tactile glove (inner glove only), glove with protective cover.**

This kind of feedback can improve spatial awareness about objects surrounding the hand to guide the motion, which helps to avoid contact before it happens. While proximity cues have been introduced to optimize pointing behavior during 3D selection tasks [3], we expect such cues are also beneficial for manipulation tasks that are driven by steering behaviors. Yet, we are unaware of work that has explored proximity cues for manipulation tasks. Our proximity feedback provides continuous, spatio-temporal audio-tactile feedback about objects surrounding the hand, independent from contact events. This feedback is coupled to object collision and friction cues that relate to the biomechanical control (contact) points, to enrich task chain-driven feedback. In our approach tactile feedback only provides indications about distance to other objects, while directional information is provided through audio. We made this choice based on the results of pilot studies, described in section 5.1. Audio extends the tactile feedback by providing sound upon impact (collision), directional and distance cues to objects around the hand (proximity), and texture cues during friction. Coupling audio to tactile cues can be beneficial as there is evidence for good multisensory integration of both, especially with regards to temporal aspects [39]. However, while audio and vibration have been shown to improve performance in 2D interfaces [12], there is surprisingly little evidence for performance improvements for 3D selection and manipulation tasks.

Our feedback mechanism differs from previous work on audio-tactile proximity-based selection assistance [3] in multiple ways. There the authors used only non-directional cues and focused solely on selection, not manipulation. Also, non-directional cues can only encode distance to a single object, which is insufficient in scenes where users can collide with multiple surfaces/objects around the hand. In contrast, our approach uses a glove-based interface developed in-house that contains a higher-density grid of vibrotactors across both sides of the hand and as such provides contact information across the full hand. Moreover, we use directional cues to elicit directional information about objects in hand proximity.

### 4.1 Tactile Glove

We developed a vibrotactile glove (see Fig. 2) whose operation range supports full arm motions. Hand pose and motion is tracked through optical methods, in our case a Leap Motion. The glove has also been used for other purposes, namely hand motion and pose guidance. In [35] we illustrated how tactile patterns can guide the

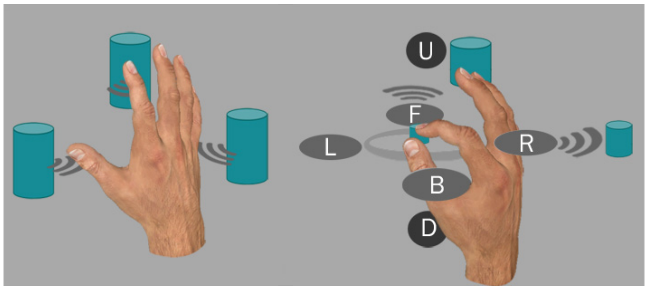

**Figure 3: Outside-in proximity cues, where audio-feedback is spatialized in the scene (Left). Inside-out proximity cues, where sound localization is tied to the hand (Right).**

user, by triggering hand pose and motion changes, for example to grasp (select) and manipulate (move) an object.

The glove is made of stretchable, comfortable-to-wear cotton. In the glove, tactors are placed at the finger tips (5 tactors), inner hand palm (7), middle phalanges (5), and the back of the hand (4), for a total of 21 tactors (Fig. 2). An outer cotton cover fits exactly over the inner glove to protect the cables and lightly press the tactors against the skin. We use 8-mm Precision Microdrive encapsulated coin vibration motors (model 308-100). In our pilot studies, we identified that tactors #19-21 lie too close to the tactor used for proximity feedback, #16. Especially during grasping, this leads to misinterpretation of cues, as tactors move closely together. Thus, we used only tactors #1-18 in our studies, to avoid confusion between collision and proximity feedback. With 18 tactors, we simulate many contact points that are associated with grasping objects (palm, fingertips) while also supporting collision feedback at the back of the hand. This is a novel feature, as back-of-the-hand feedback is generally not supported in tactile interfaces. Even though we do not cover the full hand surface with tactors, we still cover most areas and can benefit from phantom effects by interpolating between tactors, similar to [20]. The cable ends and master cable are attached at the back of the wrist through a 3D printed plate embedded in the fabric. All tactors are driven by Arduino boards. To overcome limitations in motor response caused by inertia (up to ~75 ms), we use pulse overdrive [36], which reduces latency by about 25 ms.

## 4.2 System and Implementation

The system was implemented in Unity3D V5.6, using NVidia PhysX 3.3 for collision detection. Hand tracking was performed with a Leap Motion, through the Orion SDK. We used the Uniduino plugin to control four Arduino Boards to trigger the tactors. The system ran on a graphics workstation (Core i7, 16GB RAM, NVidia 1080GTX, Windows 10) to guarantee fluid performance. During the first study, interaction was performed below a reach-in display, a 20-degree angled 32" display (Fig. 1, right). Replicating standard virtual hand metaphors [32], we only showed the hand, not the wrist or arm, using a realistic 20,000 polygon hand model from the Leap Motion SDK. The index finger and thumb were used to grab (pinch) an object. Once an object is pinched, the user receives a short tactile burst at the thumb and index fingertip. While the user holds the object, no further tactile cues are provided at these locations, to avoid habituation as well as confusion between pinch and scene collision cues.

*4.2.1 Proximity Feedback modes.* We explored two modes that combine tactile and audio feedback for proximity feedback. With outside-in feedback, each object in the scene emits signals, i.e., feedback is spatially tied to the objects in the scene. In contrast, with inside-out feedback, feedback is provided relative to the grasped object in the hand – directions are divided into zones. The hand "sends" signals out into the scene, and "receives" spatial feedback about which zones around the hand contain objects, similar to radar signals. Both modes are implemented analogous to car park assistant technologies to indicate where (direction) and how close (distance) surrounding objects are. Tactile cues are represented by vibration patterns, starting with slow and light vibrations and, as the distance to neighboring objects shortens, ending with stronger and shorter-cycle vibrations.

Vibrotactile proximity cues are provided for the closest available object collider as soon the users hand is close enough. As discussed above, we use the pulse overdrive method to quickly activate the corresponding tactor. To stably drive the motor, we then reduce the voltage via pulse width modulation (PWM) to the lowest possible amount, about 1.4V (a duty cycle of 28%). As the user is getting closer to the collider, the duty cycle is adjusted inversely proportional to the collider distance, creating the maximum vibration intensity with a duty cycle of 100% right at the object.

We use a single tactor in the palm (tactor #16 in Fig. 2) to provide vibrotactile proximity cues, and use audio to communicate the direction and distance to surrounding objects. This design decision was based on pilot studies that showed that full-hand proximity cues are difficult to separate from collision cues. Furthermore, we introduce a deliberate redundancy between tactile and auditory distance cues, as we aim to strengthen the amount of "warning" before potential object penetrations. To provide audio cues, we used the Audio Spatializer SDK of the Unity game engine. This allows to regulate the gains of the left and right ear contributions based on the distance and angle between the AudioListener and the AudioSource, to give simple directional cues.

For *outside-in* proximity feedback, each object contains a spatially localized audio source: hence, users can hear the location of the objects over the used headphones. The audio "objects" are characterized not only by their location relative to the hand, but also by volume and pitch to provide 3D spatial cues. The adjustment of volume depends on the relative distance to the hand with a linear roll-off within a specified radius. As long the hand is within the roll-off threshold, the sound starts at neutral pitch level and gets higher the closer the hand gets to an object. As it is scene-driven, we assumed this model would be beneficial for general spatial exploration tasks: the feedback provides a general indication about objects in vicinity of the hand, instead of targeting more precise cues related to a single (grasped) object.

To support *inside-out* proximity feedback, we located six audio sources around the hand that define unique directions along the coordinate system axes. If an obstacle is detected at a certain direction, the corresponding proximity sound is played with the same volume and pitch characteristics as in the selection phase. Different abstract ("humming") sounds are used for up/down proximity compared to forward/backward/left/right proximity, in order to make the cues more distinguishable. This method is similar to parking aids in cars.

Motivated by previous work [46], the pitch of a sound indicates the approximate position in the vertical direction: higher pitched sounds are perceived as originating above lower pitched sounds. As this model provides highly granular proximity cues in relation to the hand (and grasped object), we assumed that it can be beneficial for manipulation tasks in which an object is moved through a scene.

*4.2.2 Collision and Friction Feedback.* Once the user actually touches an object, we provide location-specific collision cues, based on a mapping between contact point and an adjacency list of the tactors. All motors are given an individual weighting factor (see Fig. 2) which were fine-tuned through a pilot study reflecting on the local mechanoreception properties of the skin [23]. We calculate the distance of the collision point in relation to the closest tactor on the glove. If a collision point is in between two tactors, this results in interpolation of vibration strength, similar to a mechanism described previous work [20]. Beyond the mechanoreception weighting factor, modulation of the tactor is then also affected by the distance to objects and hand velocity, resulting in a higher intensity when the collision occurs at a higher speed.

We use the Karnopp model, a stick-slip model that describes friction forces as the exceedance of the minimal tangential velocity relative to the object surface to provide friction cues [26]. Friction cues are triggered by the combination of object penetration and velocity, and are represented through both vibration and audio feedback [31]. We use the PhysX API to determine penetration and its depth. Similar to proximity, friction cues consist of localized auditory and vibrotactile feedback, while tactile cues are directly dependent on the the sound waveform that represents the material properties, similar to the method presented in [31]. For auditory friction feedback, we take the penetration depth and the velocity of the penetrating object into account. A material-conform friction sound is assigned to each object in the scene, and is faded in or out depending on penetration depth. The intensity and pattern of the vibration feedback is based on the spectrum of the played friction sound, similar to [31].

## 5 USER STUDIES

In our user studies we explored how different audio and tactile cues affect *touch* and *motion* by looking how proximity cues influence spatial awareness in a scene exploration task (RQ1, using the outside-in model) and precise object manipulation performance in a fine motor task (RQ2, with the inside-out model). All studies employed the setup described above. With consent of the users, demographic data was recorded at the start. For study 1, we only analysed subjective feedback, while for study 2 we logged task time, object collisions, penetration depth and the number of tunnel exits in between start and end position (errors). After the study, participants rated their level of agreement with several statements related to concentration, cue usefulness, perceptual intensity, and spatial awareness on a 7-point Likert scale (7 = "fully agree"). It took between 45 and 75 minutes to complete the whole study.

### 5.1 Pilot studies

We performed several pilot studies during the design and implementation process of our glove interface prior to the main ones. The first pilot aimed to verify our feedback approach, coupling proximity, collision and friction cues. Nine users (1 female, aged between 25 and 30 years) interacted with an early design of the glove. Users performed a key-lock object manipulation task, selecting a target object and moving it into another object. The objects were small and partly visually occluded. The pilot confirmed the utility of the proximity-driven approach, but identified limitations in tactile resolution and audio feedback. This informed the design of a higher-resolution glove. Based on an near-complete version of the glove, the second pilot fine-tuned feedback cues and probed study parameters for the main studies. Through multiple tests performed with 4 people we tuned the weighting factors of the tactors, with the results shown in Fig. 2. A third pilot with 6 users (one female, aged between 26 and 39) explored various design parameters of our main studies. This pilot included a tunnel task and a search task to find an opening, and was used to make final adjustments to the glove feedback mechanisms, in particular the proximity based feedback approach in the reach-in display system (Fig. 1).

### 5.2 Study 1 - Scene Exploration

In this study, we explored how the number of contact points afforded by the glove and the enabling or disabling of proximity cues affects spatial awareness in relation to hand motion constraints, i.e. hand-scene constraints, during scene exploration.

*Task.* We showed a start position and the position of an object to select, which defined the end position. We located several invisible objects (cubes) between the start (front) and end position (back), creating an environment through which the hand had to be maneuvered without touching or passing through obstacles (see Fig. 1, second image from Left). Before selecting the object, users had to explore the scene while receiving collision, proximity and friction cues, which enabled them to understand the scene structure. As the Cybertouch is currently a quasi-standard in vibrotactile gloves, the glove was either used with full resolution for collision (18 tactors) or simulating the Cybertouch II (6 tactors, one at each finger tip, one in the palm, ID 16, Fig. 2, Right). In both conditions proximity cues were only felt at the tactor at the palm of the hand. In our simulated low-resolution Cybertouch condition, collision cues were remapped to match the limited number of tactors. We compared this condition with our high-resolution tactor configuration to assess if increasing the number of tactors enables better performance. In other words, we investigated if quasi full-hand feedback instead of mainly finger-tip and palm feedback provides more benefits compared to somewhat higher technical complexity of additional tactors.

*Procedure.* The study was performed within-subjects and employed a 2 (low or high resolution feedback) x 2 (proximity feedback on / off) x 2 (different scenes) design, totaling 8 trials. All scenes had to be explored for about 1 minute each and feedback was based on the scene-driven outside-in proximity model. Participants were asked to evaluate if they could more easily judge where their hand would fit between objects depending on proximity cues (off vs. on) and the resolution of the feedback (high vs. low).

### 5.3 Study 2 - Object Manipulation

In this study, we looked into the effect of proximity cues on user performance during a manipulation tasks that involved steering the hand (with a grasped object) through a scene. We used a tunnel scene analogy as it is quite common to assess steering tasks using paths with corners [57], while it also shows resemblance to assembly tasks where a grasped object needs to be moved through space. *Task.* Users were asked to move a small object (2 cm size) through an *invisible* tunnel (from top front to lower back). Participants were instructed to move as fast as possible, while reducing collisions and penetrations with or pass-throughs of tunnel walls. In this study we always used all 18 tactors - 17 for contact information, and one for proximity. The focus of our research was on the usefulness and performance of the different feedback conditions, i.e., collision, proximity and feedback cues, during fine object manipulation. We aimed to isolate the effect of each feedback method through three blocks and also looked into potential learning effects. The tunnel contained two straight segments connected by a 90 degree corner (main axis). The “bend” was varied by changing the angle of the two connected tunnel segments (10 degrees variations from the main axis - tunnels with more angled segments were expected to be more difficult). Tunnels had a wall thickness of 1.5 cm, which was used to calculate penetration depth and pass-throughs. We only showed the start and end positions of the tunnel and the object to be selected, while the rest of the tunnel remained invisible. This forces users to focus on the tactile cues in isolation and has the additional benefit that it avoids any potential disadvantages of any given visualization method (such as depth ambiguities associated with transparency). When users exited the tunnel by more than 1 cm between start and end, users had to restart the trial. Users wore the glove (Fig. 2), while interacting underneath the reach-in display. To avoid the potential confound of external auditory cues during the user studies and to remove the effect of potential audio disturbances, we used Bose 25 headphones with active noise cancellation.

*Procedure.* This study used the object-driven inside-out proximity model. It deployed a within-subject design, and consisted of three blocks. Block 1 (*collision only*) included 9 trials, defined by the nine tunnel variants (3 variants of segment one x 3 variants of segment two). Subjects performed the task solely with collision feedback. This block implicitly also familiarized participants with the procedure. Block 2 (*collision and proximity*) employed a 9 (tunnel variants) x 2 (with and without audio proximity cues) x 2 (with and without vibration proximity cues) factorial design, totaling 36 trials. Collision feedback was always enabled. Block 3 (*collision, proximity and friction*) employed a 9 (tunnel variants) x 2 (with or without friction) factorial design, totaling 18 trials, where collision and audio-tactile proximity cues were always enabled. We split the experiment into blocks, as a straight four-factor design is statistically inadvisable. Instead, our blocks build on each other, which enables the comparison of trials with and without each cue. Between blocks participants were introduced to the next feedback condition in a training scene. As friction cues alone do not help to avoid collisions they were only presented in combination with proximity cues in the third block. It took around 35 minutes to finish this study.

## 5.4 Results

The sample for study 1 and 2 was composed of 12 right-handed persons (2 females, mean age 31.7, SD 11.11, with a range of 23–58 years). Five wore glasses or contact lenses and 7 had normal vision. The majority played video games regularly, 6 persons daily (50%), 5 weekly (41.7%) and one only monthly (8.3%). All participants volunteered and entered into a drawing (with a shopping voucher).

**Table 1: Mean ratings (standard deviations in brackets) during scene exploration, for hand-scene constraints with proximity cues ("does the hand fit through") and contact points.**

| | Feedback Resolution | | |
|---|---|---|---|
| Perceived constraints | low | high | Improvement |
| – off | 4.08 (0.90) | 4.92 (0.90) | +20.6% $^{**}$ |
| – on | 5.33 (0.88) | 6.25 (0.62) | +17.3% $^{**}$ |
| Improvement | +30.6% $^{***}$ | +27.0% $^{***}$ | |
| Perceived contact point | | | |
| – overall hand | 4.08 (0.90) | 5.33 (1.37) | +30.6% $^{*}$ |
| – fingers | 4.50 (1.24) | 5.67 (1.37) | +26.0% $^{**}$ |
| – back of hand | 3.42 (1.08) | 5.0 (1.04) | +46.2% $^{**}$ |
| – palm | 4.33 (1.37) | 4.92 (1.24) | +13.6%, n.s. |

$^{*}$ $p < .05$, $^{**}$ $p < .01$, $^{***}$ $p < .001$

*5.4.1 Study 1.* In this part of the study, participants explored a scene to gain spatial awareness of the scene structure. As this task was not performance driven, we only report on subjective ratings from the questionnaire, analyzed using paired t-tests.

Table 1 shows mean ratings and standard deviations as well as statistically significant differences. The mean level of agreement was significantly higher for high resolution than for low resolution feedback, both with proximity cues and without. Comparing the ratings for the same statement between proximity cues (off vs. on), the level of agreement was higher with proximity cues than without in both the high and low resolution feedback conditions. The point of collision could be better understood with high than with low resolution feedback on the overall hand, fingers, and the back of the hand, but not in the palm.

*5.4.2 Study 2.* For the analysis of blocks 1 to 3, we used in each case a repeated-measures ANOVA with the Greenhouse-Geisser method for correcting violations of sphericity, if necessary. Dependent variables were time to finish a trial successfully, collisions, penetration depth and errors in each block. Independent variables differed between blocks, in the first block we examined the effect of tunnel variants, in the second block the effect of tunnel variants, proximity audio and vibration cues and in the third block the additional use of friction. The effect of the factor cue on different questionnaire ratings for block 2 was examined using a one-way repeated measures ANOVA. Post-hoc comparisons were SIDAK corrected. For block 3, paired t-tests were used to compare questionnaire ratings for trials with and without friction cues. All tests used an alpha of .05. Below, we only report on the main results.

Time to finish a trial increased between blocks (31.06 s, SD=23.4 for block 1, 36.65 s, SD=27.52 for block 2, 40.15 s, SD=28.64 for block 3). In block 1 (collision cues only) there was no effect of the tunnel variants in terms of collisions, penetration depth or errors, except that there was an effect on time, F(8,88) = 2.16, p = .038, $\eta^2 = .16$. As expected, tunnels with angled segments took longer.

In block 2 we analyzed collision and proximity cues. The time required to pass tunnels was not affected by the tunnel variant, and

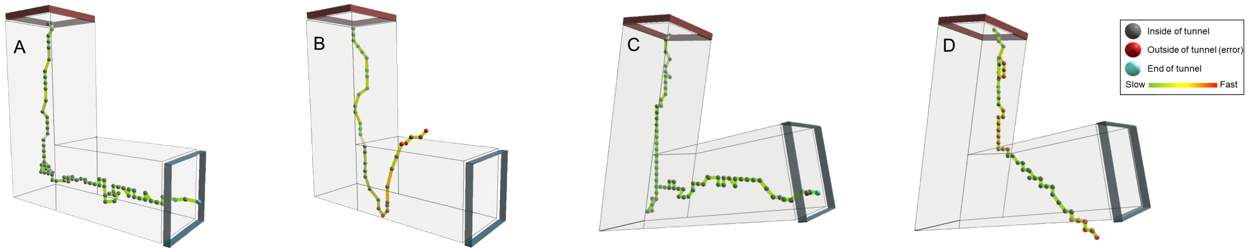

**Figure 4: Example paths from Study 2. The first tunnel is simple, with a 90° bend (A & B). The second variant is moderately difficult, with a 70° turn (C & D). Tunnel walls were not visible to participants in the studies.**

was also not influenced by cues. Yet, the tunnel variants significantly influenced the number of collisions, F(8,88) = 2.64, p = .012, $\eta^2$ = .19. Most tunnels produced a limited range of collisions, 3.51 (SD = 3.25) to 5.71 (SD = 3.57), except for the most complex one that produced 7.50, (SD = 6.57). For proximity cues we observed that most collisions occurred when both cues were off and fewest when only audio cues were on (Table 2 shows mean values and significances).

**Table 2: Study 2, block 2: Mean performance values depending on proximity cues and % change against baseline. *Prox* stands for proximity, *A* for audio, *V* for vibration**

| | Collisions | Penetration depth | Errors |
|---|---|---|---|
| *Collision* (baseline) | 6.17 | 0.145 | 1.56 |
| *Prox − A* only | 4.3 *<br>(-30.3%) | 0.125 **<br>(-13.8%) | 0.68 **<br>(-56.4%) |
| *Prox − V* only | 4.94 *<br>(-19.9%) | 0.142 n.s.<br>(-2.1%) | 1.13 n.s.<br>(-27.6%) |
| *Prox − A + V* | 4.56 n.s.<br>(-24.6%) | 0.113 **<br>(-22.1%) | 0.9 n.s.<br>(-42.3%) |

n.s. not significant, * $p < .05$, ** $p < .01$

Audio and vibration proximity cues showed no main effect on the number of collisions, but there was a tendency to an interaction effect of proximity cues, F(1,11) = 4.76, p = .052, $\eta^2$ = .30 (see Fig. 5). Post-hoc comparisons revealed that audio or vibration proximity cues alone significantly affected the number of collisions when the other proximity cue was turned off (p < .05). Furthermore, mean penetration depth was significantly smaller with audio cues (Table 2, F(1,11) = 14.57, p = .003, $\eta^2$ = .57). Penetration depth was also influenced by the tunnel variant (F(4.50,49.48) = 4.34, p = .003, $\eta^2$ = .28) – again, the most complex one lead to the largest penetration depth (M = 0.155, SD = 0.036), Regarding errors there was a tendency to an interaction effect of audio and vibration proximity cues, F(1,11)= 4.55, p = .056, $\eta^2$ = .29 see Fig. 5. When vibration proximity cues were turned off, audio proximity cues significantly influenced the number of errors as less errors occurred with audio proximity cues than without (Table 1, p = .035). The presence of vibration cues did not significantly reduce the number of errors when audio was turned off (p = .093).

Block 3 focused on collision, proximity and friction cues. There was a significant effect of tunnel variant on the number of collisions (F(8,88) = 4.38, p < .001, $\eta^2$ = .29), but no effect on time, mean penetration depth and errors. Again the most complex tunnel stood out, with the most collisions (M = 8.17, SD = 6.24). Friction cues did not affect any of the dependent variables and there was also no interaction effect of tunnel variant and friction.

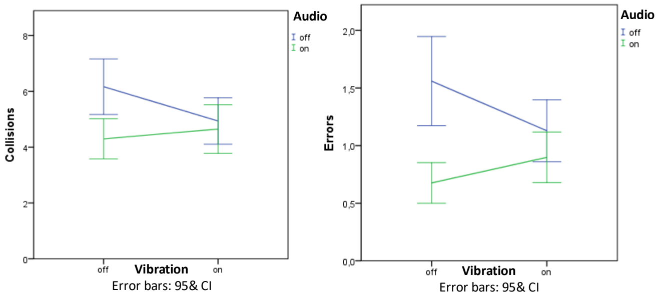

**Figure 5: The effect of audio and vibration proximity cues on collisions and errors.**

## 5.5 Path analysis

To better understand participant performance during the trials, we sampled the dataset by selecting best and worse trials from different tunnel conditions (easy and more difficult ones, as defined by the variable angle between both tunnel segments). Here, we present the most relevant examples of this process to exemplify path behavior. Fig. 4A & B show examples of an easy task (90° bend) in visual comparison to a more challenging one (70° turn, Fig. 4C & D). With all activated proximity cues (collision, proximity and friction cues, Fig. 4A & C) participants found it easier to stay within the tunnel, while this was harder when only collision cues were present (Fig. 4B & D). In the latter cases the path shows only a partial run until the first error occurred (which required a restart of the trial). Samples and measurements taken at various points along the path of the examples paths show that proximity cues can help the user to move the object closer along the ideal path for both easy task (M = 0.69, Fig. 4A) and difficult task (M = 0.71, Fig. 4C). In contrast, however, without proximity feedback, the distance to the ideal path increased drastically for the simple task (M = 0.86, Fig. 4B), as well as for the difficult task (M = 1.22, Fig. 4D). This resulted in a higher error rate, through participants (unintentionally) leaving the tunnel.

Manipulation behavior is different from selection. Selection is a pointing task that exhibits a ballistic, fast phase before a corrective, slower motion phase. In contrast, a manipulation is a steering task in which motion velocity is far more equalized [42, 57]. As such, manipulation performance – and difficulty – is affected by the steering law, instead of Fitts's law [42]. Like Fitts's law, steering difficulty is defined by path width and curvature, yet is linear instead of logarithmic. The absence of velocity difference due to ballistic and corrective motions hand motions can be clearly seen in our examples. While velocity varies from about 14.14mm/s to 67.12mm/s in the shown samples, fast movements are only performed rarely, and not in patterns that conform to rapid aimed pointing movements. Of course, steering still exhibits corrective motions, as can be seen for example in Fig. 4B at the lower end of the path. What is also striking is the behavior of steering through corners: the path does not necessarily adhere to the shortest path (hence, cutting the corner), rather the ideal path is defined by staying clear of the corners [42], even though Fig 4D shows this is not always successful. This is somewhat in contrast to behavior in 2D interfaces, as noted in [42], where corners tended to get cut. We assume that in our case, cutting was avoided as proximity cues encourages the user to stay away from surrounding objects and thus also corners.

**Table 3: Mean level of agreement on 7-point Likert items and standard deviations for cue usefulness in study 2, block 2 & 3. *Prox* stands for proximity, *A* for audio, *V* for vibration, *Fric* for friction.**

| | Performed faster | Performed more precisely | Understood the tunnel shape better | Reacted more quickly |
|---|---|---|---|---|
| $Collision$ | 4.92 (1.78) | 5.17 (1.27) | 4.83 (1.53) | 5.58 (1.73) |
| $Prox - A$ | 5.58 (1.38) | 5.58 (1.83) | 5.75 (1.49) | 5.92 (1.08) |
| $Prox - V$ | 5.33 (1.16) | 5.33 (1.44) | 5.08 (1.44) | 5.17 (1.12) |
| $Prox - A + V$ | 5.42 (1.62) | 5.67 (1.78) | 5.75 (1.49) | 5.75 (1.55) |
| $Prox - A + V$ | 4.42 (1.08) | 4.67 (1.23) | 4.92 (1.31) | 4.83 (1.40) |
| $\ldots + Fric$ | 5.25 (1.22) | 5.67 (1.23) | 6.0 (1.35) | 5.83 (1.27) |
| Improvement | +18.8% * | +21.4% * | +22% * | +20.7% * |

$^{*}\ p < .05$

## 5.6 Subjective Feedback

Questionnaire ratings indicated that all cues facilitated to perform the task faster and more precisely, aided understanding of the tunnel shape, and made movement adjustments easier (Table 3). However, there was no significant difference between cue ratings. Interestingly, participants thought they performed the task faster (t(11) = -2.59, p = .025), more precisely (t(11) = -2.71, p = .02), understood the shape of the tunnel better (t(11) = -2.86, p = .015), and reacted more quickly to adjust the object movement in the scene (t(11) = -2.45, p = .032) while using friction. In the open comments it was also striking that half of the participants reported that it was easier to focus on a single proximity cue at any given time. Some users stated they experienced a limited form of information overload when both proximity cues were activated simultaneously, which distracted them. Finally, we also evaluated the overall usability, comfort and fatigue in the questionnaire (see Table 4). Most ratings were positive to very positive, though tracking errors and cabling issues were noted. As the experiment took some time we were particularly interested in user fatigue. Fortunately participants rather disagreed that they got tired while wearing the glove interface.

**Table 4: Mean level of agreement with comfort and usability statements on 7-point Likert items and standard deviations**

| Statement | Mean Rating (SD) |
|---|---|
| Sitting comfort | 5.33 (1.14) |
| Glove wearing comfort | 6.42 (0.67) |
| No disruption through the cable | 3.25 (1.71) |
| Match of virtual to real hand | 5.25 (1.14) |
| Hand tracking problems | 4.41 (1.78) |
| Ease of learning the system | 5.5 (1.24) |
| Ease of using the system | 5.58 (1.17) |
| Expected improvement through exercise | 6 (0.74) |
| Getting tired wearing the glove interface | 3.25 (1.49) |

## 5.7 Discussion

In our studies, we investigated the effect of proximity cues for hand touch and motion associated with scene exploration and manipulation actions. Here, we discuss our main findings.

*RQ1. Do scene-driven proximity cues improve spatial awareness while exploring the scene?*

Overall, our scene exploration study provides positive indications about the usage of scene-driven outside-in proximity cues to enhance spatial awareness. It also indicates a positive effect of increasing the number of tactors, as both the awareness of hand-scene constraints and contact (touch) points across the hand improved. The performance improvements provide a positive indication for higher numbers of tactors in novel glove-based or other types of full-hand interfaces. With our high-density tactor design, the localization of contact points across the hand improved about 30% in comparison to a Cybertouch-like configuration. It is also interesting to contrast our results to the hand-palm system TacTool that uses six vibration motors[45]. There, directionality (mainly of collision cues) was not always easily identified, whereas in our system, the simulated contact point was always well differentiated. While a contact point alone does not indicate an exact impact vector, it enables at least an identification of the general impact direction. Potential explanations for our different finding include the different locations and numbers of tactors, as well as a different hand posture. Finally, as the inside-out model partitions surroundings into zones irrespective of the amount of objects, we assume that our approach is resilient towards increasing object density in a scene, but have not yet verified this.

*RQ2. Can hand-driven proximity cues avoid unwanted object penetration or even **touching** proximate objects during manipulation tasks?*

In our manipulation task, we showed that audio-tactile proximity cues provided by the object-driven inside-out model significantly reduced the number of object collisions up to 30.3% and errors (object pass-throughs) up to 56.4%. With touch cues users thought they could perform faster (18.8%), more precise (21.4%), and adjust hand motion quicker (20.7%). Interestingly, audio cues alone also produced surprisingly good results, which is a useful finding as it potentially frees up vibrotaction for purposes other than proximity feedback. As fewer errors were made, we assume that proximity cues can enhance motor learning. Also, as haptic feedback plays a key role in assembly procedures [43], additional cues may not only optimize motion, but also hand poses. While we only indirectly steer hand poses in this work, explicit pose guidance might a worthwhile extension [6]. Interestingly, our results somewhat contradict previous findings that identified bimodal feedback to be less beneficial in terms of throughput [3]. While we cannot calculate throughput for the steering task users performed, it would be interesting to investigate the measure on simpler tasks with our proximity models. Also, while we currently have a uniform tunnel width, it will be interesting to contrast our results to other tunnel widths in future work. Furthermore, users noted in their subjective feedback that single cues were, not entirely unexpected, easier to focus on than coupled cues. However, while cognitive load may pose an issue, it is not uncommon for multimodal interfaces to increase load [56]. In this respect, it is worth to mention related work [49] that has looked into bimodal (audio-tactile) and unimodal (tactile) feedback in touch related tasks. Results revealed a significant performance increase only *after* a switch from bimodal to unimodal feedback. The authors concluded that the release of bimodal identification (from audio-tactile to tactile-only) was beneficial. However, this benefit was not achieved in the reverse order. The interplay between modalities also gives rise to potential cross-modal effects. Previous work in the field of object processing using neuroimaging methods [27] has shown multisensory interactions at different hierarchical stages of auditory and haptic object processing. However, it remains to be seen how audio and tactile cues are merged for other tasks in the brain and how this may affect performance.

Overall, through our fully directional feedback, we extend previous findings on single-point, non-directional proximity feedback [3] that elicit constraints on dimensionality. We confirm that directional feedback can improve performance, in particular through a reduction of errors. We also improve on previous work by investigating fully three-dimensional environments. In this context, it would be interesting to assess performance differences between non-directional and directional feedback in the future, also for selection tasks, while also looking more closely at potential learning effects. While we focused on the usefulness of proximity feedback in manipulation tasks, we expect our inside-out feedback to also have a positive effect on selection tasks. Another open area is the trade-off and switching between outside-in and inside-out proximity feedback models based on the usage mode (selection versus manipulation versus exploration). Such switching has the potential to confuse users and thus necessitates further study.

Similar to Ariza et al. [3], we studied the feedback methods in the absence of additional visual feedback in this work. This poses the question how our methods can be used in combination with visual feedback, and what dependencies any given visualization technique introduces in a real usage scenario. Naturally, information about objects around the hand is usually communicated over the general visual representation of the rendered objects, as will be the case during, e.g., learning assembly procedures. Yet performance may be affected by visual ambiguities. While visual and haptic stimuli integration theories [13] underline the potential of a close coupling of visual and non-visual proximity cues, ambiguities may still affect performance. Researchers have looked into reducing such ambiguities, for example through transparency or cut-away visualizations, where spatial understanding may vary [14]. Another approach to address ambiguities might be to provide hand co-located feedback, where first attempts have been presented previously, e.g., [44]. For example, portions of the hand could be color coded based on their level of penetration into surrounding objects. Hence, we are considering to verify performance of our methods in combination with visual feedback in the future, using both standard or optimized visualization methods.

## 6 CONCLUSION

In this work, we explored new approaches to provide proximity cues about objects around the hand to improve hand motor planning and action coordination during 3D interaction. We investigated the usefulness of two feedback models, outside-in and inside-out, for spatial exploration and manipulation. Such guidance can be highly useful for 3D interaction in applications that suffer from, e.g., visual occlusions. We showed that proximity cues can significantly improve spatial awareness and performance by reducing the number of object collisions and errors, addressing some of the main problems associated with motor planning and action coordination in scenes with visual constraints, which also reduced inadvertent pass-through behaviors. As such, our results can inform the development of novel 3D manipulation techniques that use tactile feedback to improve interaction performance. A logical next step require integrating our new methods into actual 3D selection and manipulation techniques, while also studying the interplay with different forms of visualization (e.g., [51]) in application scenarios. In due course, the usage and usefulness of two gloves with audio-tactile cues is an interesting venue of future work, e..g, to see if audio cues can be mapped to a certain hand. Furthermore, we currently focused only on haptic feedback to eliminate potential effects of any given visualization method, such as depth perception issues caused by transparency. Finally, we are looking at creating a wireless version of the glove and to improve tracking further, e.g., by using multiple Leap Motion cameras [21].

## 7 ACKNOWLEDGEMENTS

This work was partially supported by the Deutsche Forschungsgemeinschaft (KR 4521/2-1) and the Volkswagen Foundation through a Lichtenbergprofessorship.

## REFERENCES

[1] C. Afonso and S. Beckhaus. 2011. How to Not Hit a Virtual Wall: Aural Spatial Awareness for Collision Avoidance in Virtual Environments. In *Proceedings of the 6th Audio Mostly Conference: A Conference on Interaction with Sound (AM '11).* ACM, 101–108.

[2] N. Ariza, P. Lubos, F. Steinicke, and G. Bruder. 2015. Ring-shaped Haptic Device with Vibrotactile Feedback Patterns to Support Natural Spatial Interaction. In *ICAT - EGVE '15 Proceedings of the 25th International Conference on Artificial Reality and Telexistence and 20th Eurographics Symposium on Virtual Environments.*

[3] O. Ariza, G. Bruder, N. Katzakis, and F. Steinicke. 2018. Analysis of Proximity-Based Multimodal Feedback for 3D Selection in Immersive Virtual Environments. In *Proceedings of IEEE Virtual Reality (VR).*

[4] S. Beckhaus, F. Ritter, and T. Strothotte. 2000. CubicalPath-dynamic potential fields for guided exploration in virtual environments. In *Proceedings the Eighth Pacific Conference on Computer Graphics and Applications.* 387–459.

[5] H. Benko, C. Holz, M. Sinclair, and E. Ofek. 2016. NormalTouch and TextureTouch: High-fidelity 3D Haptic Shape Rendering on Handheld Virtual Reality Controllers. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology (UIST '16).* ACM, 717–728.

[6] A. Bloomfield, Y. Deng, J. Wampler, P. Rondot, D. Harth, M. McManus, and N. Badler. 2003. A taxonomy and comparison of haptic actions for disassembly tasks. In *Virtual Reality, 2003. Proceedings. IEEE.* IEEE, 225–231.

[7] M. Bouzit, G. Burdea, G. Popescu, and R. Boian. 2002. The Rutgers Master II-new design force-feedback glove. *IEEE/ASME Transactions on mechatronics* 7, 2 (2002), 256–263.

[8] G. Burdea. 1996. *Force and Touch Feedback for Virtual Reality.* John Wiley & Sons, Inc.

[9] L. Chan, R. Liang, M. Tsai, C. Cheng, K.and Su, M. Chen, W. Cheng, and B. Chen. 2013. FingerPad: Private and Subtle Interaction Using Fingertips. In *Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology (UIST '13).* ACM, 255–260.

[10] E. Chancey, J. Brill, A. Sitz, U. Schmuntzsch, and J. Bliss. 2014. Vibrotactile Stimuli Parameters on Detection Reaction Times. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 58, 1 (2014), 1701–1705.

[11] W. Chang, W. Hwang, and Y. Ji. 2011. Haptic seat interfaces for driver information and warning systems. *International Journal of Human-Computer Interaction* 27, 12 (2011), 1119–1132.

[12] A. Cockburn and S. Brewster. 2005. Multimodal feedback for the acquisition of small targets. *Ergonomics* 48, 9 (2005), 1129–1150.

[13] M. Ernst and M. Banks. 2002. Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* 415, 6870 (2002), 429.

[14] A. Kunert C. Andujar F. Argelaguet, A. Kulik and B. Froehlich. 2011. See-through techniques for referential awareness in collaborative virtual reality. *International Journal of Human-Computer Studies* 69, 6 (2011), 387–400.

[15] P. Gallotti, A. Raposo, and L. Soares. 2011. v-Glove: A 3D Virtual Touch Interface. In *2011 XIII Symposium on Virtual Reality.* 242–251.

[16] U. Gollner, T. Bieling, and G. Joost. 2012. Mobile Lorm Glove: Introducing a Communication Device for Deaf-blind People. In *Proceedings of the Sixth International Conference on Tangible, Embedded and Embodied Interaction (TEI '12).* ACM, 127–130.

[17] J. Hartcher-O'Brien, M. Auvray, and V. Hayward. 2015. Perception of distance-to-obstacle through time-delayed tactile feedback. In *2015 IEEE World Haptics Conference (WHC).* 7–12.

[18] C. Hatzfeld and T.A. Kern. 2014. *Engineering Haptic Devices: A Beginner's Guide.* Springer London.

[19] B. Holbert. 2007. *Enhanced Targeting in a Haptic User Interface for the Physically Disabled Using a Force Feedback Mouse.* Ph.D. Dissertation. Advisor(s) Huber, Manfred. AAI3277666.

[20] A. Israr and I. Poupyrev. 2011. Tactile Brush: Drawing on Skin with a Tactile Grid Display. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11).* ACM, 2019–2028.

[21] H. Jin, Q. Chen, Z. Chen, Y. Hu, and J. Zhang. 2016. Multi-LeapMotion sensor based demonstration for robotic refine tabletop object manipulation task. *CAAI Transactions on Intelligence Technology* 1, 1 (2016), 104 – 113.

[22] R. Johansson and J. Flanagan. 2009. Coding and use of tactile signals from the fingertips in object manipulation tasks. *Nature reviews. Neuroscience* 10, 5 (2009), 345.

[23] R. Johansson and A. Vallbo. 1979. Tactile sensibility in the human hand: relative and absolute densities of four types of mechanoreceptive units in glabrous skin. *The Journal of physiology* 286, 1 (1979), 283–300.

[24] K. Johnson and S. Hsiao. 1992. Neural mechanisms of tactual form and texture perception. *Annual review of neuroscience* 15, 1 (1992), 227–250.

[25] K. A Kaczmarek, J. Webster, P. Bach-y Rita, and W. Tompkins. 1991. Electrotactile and vibrotactile displays for sensory substitution systems. *IEEE Transactions on Biomedical Engineering* 38, 1 (1991), 1–16.

[26] D. Karnopp. 1985. Computer simulation of stick-slip friction in mechanical dynamic systems. *J. Dyn. Syst. Meas. Control.* 107, 1 (1985), 100–103.

[27] T. Kassuba, M. Menz, B. RÃűder, and H. Siebner. 2013. Multisensory Interactions between Auditory and Haptic Object Recognition. *Cerebral Cortex* 23, 5 (2013), 1097–1107.

[28] K. Kozak, J. Pohl, W. Birk, J. Greenberg, B. Artz, M. Blommer, L. Cathey, and R. Curry. 2006. Evaluation of Lane Departure Warnings for Drowsy Drivers. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 50, 22 (2006), 2400–2404.

[29] E. Kruijff, A. Marquardt, C. Trepkowski, R. Lindeman, A. Hinkenjann, J. Maiero, and B. Riecke. 2016. On Your Feet!: Enhancing Vection in Leaning-Based Interfaces Through Multisensory Stimuli. In *Proceedings of the 2016 Symposium on Spatial User Interaction (SUI '16).* ACM, New York, NY, USA, 149–158.

[30] E. Kruijff, A. Marquardt, C. Trepkowski, J. Schild, and A. Hinkenjann. 2017. Designed Emotions: Challenges and Potential Methodologies for Improving Multisensory Cues to Enhance User Engagement in Immersive Systems. *Vis. Comput.* 33, 4 (April 2017), 471–488.

[31] E. Kruijff, G. Wesche, K. Riege, G. Goebbels, M. Kunstman, and D. Schmalstieg. 2006. Tactylus, a Pen-input Device Exploring Audiotactile Sensory Binding. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology (VRST '06).* ACM, 312–315.

[32] J.J. LaViola, E. Kruijff, R.P. McMahan, D. Bowman, and I.P. Poupyrev. 2017. *3D User Interfaces: Theory and Practice.* Pearson Education.

[33] R. Lindeman, R. Page, Y. Yanagida, and J. Sibert. 2004. Towards full-body haptic feedback: the design and deployment of a spatialized vibrotactile feedback system. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology, VRST 2004.* 146–149.

[34] L. Liu and R. van Liere. 2009. Designing 3D Selection Techniques Using Ballistic and Corrective Movements. In *Proceedings of the 15th Joint Virtual Reality Eurographics Conference on Virtual Environments (JVRC'09).* Eurographics Association, 1–8.

[35] A. Marquardt, J. Maiero, E. Kruijff, C. Trepkowski, A. Schwandt, A. Hinkenjann, J. Schoening, and W. Stuerzlinger. 2018. Tactile Hand Motion and Pose Guidance for 3D Interaction. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology (VRST '18).* ACM.

[36] J. Martinez, A. Garcia, M. Oliver, J. P. Molina, and P. Gonzalez. 2016. Identifying Virtual 3D Geometric Shapes with a Vibrotactile Glove. *IEEE Computer Graphics and Applications* 36, 1 (2016), 42–51.

[37] V. Mateevitsi, B. Haggadone, J. Leigh, B. Kunzer, and R. Kenyon. 2013. Sensing the environment through SpiderSense. In *Proceedings of the 4th augmented human international conference.* ACM, 51–57.

[38] M. Mine, F. Brooks, Jr., and Carlo H. Sequin. 1997. Moving Objects in Space: Exploiting Proprioception in Virtual-environment Interaction. In *Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH '97).* ACM Press/Addison-Wesley Publishing Co., 19–26.

[39] V. Occelli, Charles Spence, and Massimiliano Zampini. 2011. Audiotactile interactions in temporal perception. *Psychonomic Bulletin & Review* 18, 3 (01 Jun 2011), 429–454.

[40] T. Oron-Gilad, J. L. Downs, R. D. Gilson, and P. A. Hancock. 2007. Vibrotactile Guidance Cues for Target Acquisition. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* 37, 5 (2007), 993–1004.

[41] D. Pai. 2005. Multisensory interaction: Real and virtual. In *Robotics Research. The Eleventh International Symposium.* Springer, 489–498.

[42] R. Pastel. 2006. Measuring the Difficulty of Steering Through Corners. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '06).* ACM, 1087–1096.

[43] B. Petzold, M. Zaeh, B. Faerber, B. Deml, H. Egermeier, J. Schilp, and S. Clarke. 2004. A Study on Visual, Auditory, and Haptic Feedback for Assembly Tasks. *Presence: Teleoper. Virtual Environ.* 13, 1 (Feb. 2004), 16–21.

[44] M. Prachyabrued and C. W. Borst. 2014. Visual feedback for virtual grasping. In *2014 IEEE Symposium on 3D User Interfaces (3DUI).* 19–26.

[45] H. Regenbrecht, J. Hauber, R. Schoenfelder, and A. Maegerlein. 2005. 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 South East Asia (GRAPHITE '05).* ACM, 381–387.

[46] S. Roffler and R. Butler. 1968. Localization of tonal stimuli in the vertical plane. *The Journal of the Acoustical Society of America* 43, 6 (1968), 1260–1266.

[47] R.A. Schmidt and C.A. Wrisberg. 2004. *Motor Learning and Performance.* Human Kinetics.

[48] C. Seim, N. Doering, Y. Zhang, W. Stuerzlinger, and T. Starner. 2017. Passive Haptic Training to Improve Speed and Performance on a Keypad. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 1, 3, Article 100 (Sept. 2017), 13 pages.

[49] S.Hazenberg and R. van Lier. 2016. Touching and Hearing Unseen Objects: Multisensory Effects on Scene Recognition. *i-Perception* 7, 4 (2016).

[50] C. Spence and S. Squire. 2003. Multisensory integration: maintaining the perception of synchrony. *Current Biology* 13, 13 (2003), R519–R521.

[51] J. Sreng, A. Lecuyer, C. Megard, and C. Andriot. 2006. Using Visual Cues of Contact to Improve Interactive Manipulation of Virtual Objects in Industrial Assembly/Maintenance Simulations. *IEEE Transactions on Visualization and Computer Graphics* 12, 5 (2006), 1013–1020.

[52] H. Uchiyama, M. Covington, and W. Potter. 2008. Vibrotactile Glove guidance for semi-autonomous wheelchair operations. In *Proceedings of the 46th Annual Southeast Regional Conference on XX.* ACM, 336–339.

[53] N. van Atteveldt, M. Murray, G. Thut, and C. Schroeder. 2014. Multisensory Integration: Flexible Use of General Operations. *Neuron* 81, 6 (2014), 1240 – 1253.

[54] R. Van der Linde, P. Lammertse, E. Frederiksen, and B. Ruiter. 2002. The HapticMaster, a new high-performance haptic interface. In *Proc. Eurohaptics.* 1–5.

[55] S. Vishniakou, B. Lewis, X. Niu, A. Kargar, K. Sun, M. Kalajian, N. Park, M. Yang, Y. Jing, P. Brochu, et al. 2013. Tactile Feedback Display with Spatial and Temporal Resolutions. *Scientific reports* 3 (2013), 2521.

[56] H. Vitense, J. Jacko, and V. Emery. 2002. Multimodal Feedback: Establishing a Performance Baseline for Improved Access by Individuals with Visual Impairments. In *Proceedings of the Fifth International ACM Conference on Assistive Technologies (Assets '02).* ACM, 49–56.

[57] S. Yamanaka, W. Stuerzlinger, and H. Miyashita. 2018. Steering Through Successive Objects. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18).* ACM, Article 603, 13 pages.