# FEETICHE: FEET Input for Contactless Hand gEsture Interaction


Daniel Simões Lopes
INESC-ID / Instituto Superior Técnico, Universidade de Lisboa
daniel.lopes@inesc-id.pt

Filipe Relvas
INESC-ID / Instituto Superior Técnico, Universidade de Lisboa

Soraia Paulo
INESC-ID / Instituto Superior Técnico, Universidade de Lisboa

Yosra Rekik
Université Polytechnique Hauts-de-France, CNRS UMR 8201, LAMIH

Laurent Grisoni
Université de Lille, CNRS UMR 9189, CRIStAL

Joaquim Jorge
INESC-ID / Instituto Superior Técnico, Universidade de Lisboa
jorgej@acm.org


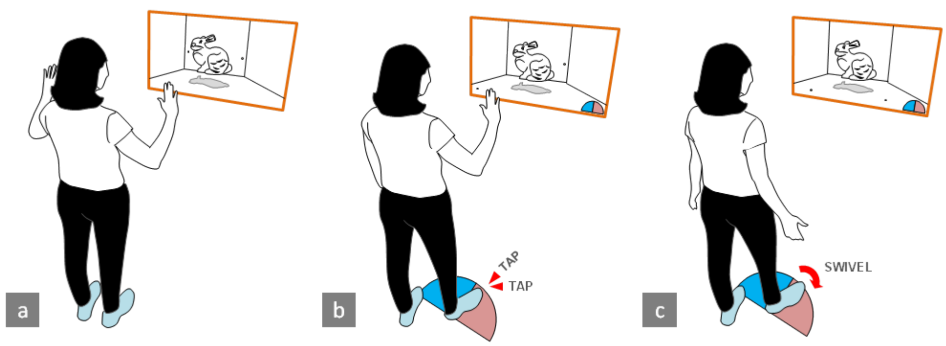


**Figure 1: Contactless manipulation of 3D objects using (a) hand and feet input to support (b) command selection and (c) geometric transformations.**


## ABSTRACT

Foot input has been proposed to support hand gestures in many interactive contexts, however, little attention has been given contactless 3D object manipulation. This is important since many applications, namely sterile surgical theaters require contactless operation. However, relying solely on hand gestures makes it difficult to specify precise interactions since hand movements are difficult to segment into command and interaction modes. The unfortunate results range from unintended activations, to noisy interactions and misrecognized commands. In this paper, we present *FEETICHE* a novel set of multi-modal interactions combining hand and foot input for supporting contactless 3D manipulation tasks, while standing in front of large displays driven by foot tapping and heel rotation. We use depth sensing cameras to capture both hand and feet gestures, and developed a simple yet robust motion capture method to track dominant foot input. Through two experiments, we assess how well foot gestures support mode switching and how this frees the hands to perform accurate manipulation tasks. Results indicate that users effectively rely on foot gestures to improve mode switching and reveal improved accuracy on both rotation and translation tasks.




*VRCAI '19, November 14–16, 2019, Brisbane, QLD, Australia*



ACM ISBN 978-1-4503-7002-8/19/11…$15.00
https://doi.org/10.1145/3359997.3365704

## CCS CONCEPTS

• **Human-centered computing → Gestural input**; *Pointing*.

## KEYWORDS

Foot interaction; hand gestures; selection; 3D manipulation; large screens; tapping; heel rotation.

**ACM Reference Format:**
Daniel Simões Lopes, Filipe Relvas, Soraia Paulo, Yosra Rekik, Laurent Grisoni, and Joaquim Jorge. 2019. FEETICHE: FEET Input for Contactless Hand gEsture Interaction. In *The 17th International Conference on Virtual-Reality Continuum and its Applications in Industry (VRCAI '19), November 14–16, 2019, Brisbane, QLD, Australia.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3359997.3365704


## 1 INTRODUCTION

Low-cost contactless 3D virtual object manipulations have been made possible via depth sensing cameras such as the Microsoft *Kinect* sensor. A typical setup involves people standing in front of large displays controlling 3D content with their bare hands. Such contactless interfaces are particularly useful and provide interesting interaction techniques for distant viewing and distant content manipulation applications, namely 3D modeling [Vinayak et al. 2013], shape externalization for object retrieval [Holz and Wilson 2011], geospatial navigation [Göbel et al. 2013; Schöning et al. 2009] and stationary Augmented Reality applications for surgical navigation [Lopes et al. 2017] where sterile environments become mandatory and preclude using devices such as keyboards or mice. Even though hand gestures are considered the dominant input modality of contactless interfaces, current hardware and software limitations make hand position tracking and gesture recognition prone to errors: (i) depth sensing cameras fail to recognize hand position whenever hands are positioned too close to each other or when they are positioned near to the body and/or close to the medial axis of the body; and (ii) unintended actions often occur at the beginning and ending of a mode switching command, that introduces small but undesired motion noise, requiring operators to alternate back and forth between activation and manipulation to acquire the desired target. These errors may be disruptive as they cause interruptions that slow down and break the 3D manipulation workflow and affect the precision of the desired result.

To overcome these limitations, better ways to operate contactless interfaces are necessary. One interesting approach to contactless interaction is to use foot input [Crossan et al. 2010; Paelke et al. 2004; Scott et al. 2010; Velloso et al. 2015]. By delegating partial control to the feet, people can improve their dexterity by assisting tasks normally performed by hand [Bowman et al. 2004]. In the context of contactless interfaces for 3D manipulation where a user is standing in front of a large display, we can observe that meaningful feet movement are mostly bi-dimensional because both feet are very important to ensure body stability, while precise 3D feet movement is bound to demand high expertise and quickly generates strong muscular fatigue. Furthermore, and to the best of our knowledge, there seems to exist a literature gap regarding feet in contactless interaction , since manipulation of 3D virtual objects, while standing and without resorting to wearable or foot devices, has not yet caught the community's attention [Velloso et al. 2015].

In this paper, we present *FEETICHE* a novel set of multi-modal interaction techniques that combine hand and foot input to support contactless 3D manipulation tasks. Our approach resorts to a simple and comfortable foot gesture vocabulary, since feet gestures which involve rotating the heel and lifting the toe feel more comfortable [Scott et al. 2010]. More specifically, we find that feet input can assist hands in clutching tasks, and feet gestures are precise enough to perform 1D translations and 1D rotations. Given environmental constraints (*e.g.*,sterile settings) we approach contactless interaction using neither any wearable or external devices to assist 3D manipulation-related tasks nor do we use hand menus or speech commands. Besides traditional mid-air hand gestures, we also consider foot double tapping and heel rotation. In addition, since heel rotation is a foot movement suitable for single axis transformations, all manipulations are constrained to a coordinate axis. In order to promote easy adoption, another important restriction is that the resulting interactive systems should be cost-effective (below $500), portable, and easily installed in workspaces as plug-and-play contactless interfaces are welcome in many application scenarios (e.g., surgical navigation systems).

Our contributions are directed at contactless 3D manipulation through the design of novel feet interaction techniques to support hand input in large display settings, specifically those where users perform 3D input gestures with their bare hands without handheld peripherals. Because the 3D content is handled using 2D interaction schemes, we focus our study on a subset of the standard 3D isometric transformations: we do not take into account the scale operation, and also reduce on the subset of translations that do not modify z coordinates, which is known to be less efficiently handled using standard approaches [Martinet et al. 2010] in the absence of extra depth cues. To better understand the benefits and drawbacks of such feet interfaces, we performed two formal evaluations to compare "hands only" to "hands and feet" interaction techniques for rotation and translation tasks. Our work is thus the first to propose a contactless multimodal interface combining hands and feet for 3D manipulation tasks. Besides the design of novel hand-feet interaction techniques and a proper discussion of the user's feedback, this paper also contributes an inexpensive sensor tracking method to detect foot states and estimate 3D foot position. Performance metrics and user satisfaction suggest that foot gestures support and even complement hand input in terms of clutching and 3D accuracy. In the remainder of this paper we survey the related work, describe our approach, and detail and discuss two separate experiments designed to assess the effectiveness of our interaction techniques. Finally, we conclude on our results and highlight possible future endeavors.

## 2 RELATED WORK

Many interaction paradigms, metaphors and techniques to handle 3D objects in virtual space have been proposed [Argelaguet and Andujar 2013; Bowman et al. 2004; Hinckley et al. 1994]. Since our work focuses on contactless hand- and foot- interactive systems, this section surveys interaction techniques that support bi-manual interaction and/or feet-based gestures to manipulate digital content.

### 2.1 Hand and Feet Input in HCI

Hand interaction using depth cameras is a well established input modality to unobtrusively control 3D user interfaces [Argelaguet and Andujar 2013]. Multi-touch hand gestures, also, have been extensively explored for this purpose [Argelaguet and Andujar 2013; Bowman et al. 2004]. When compared to hand interaction, much less attention has been given on how feet can achieve similar or improved interaction accomplishments [Velloso et al. 2015]. Nevertheless, since the 1980s that a number of interaction techniques using thigh, leg and, more importantly, foot input have been proposed for cursor positioning [Pearson and Weiser 1986, 1988], selecting menu items [Scott et al. 2010], ambient awareness [Rovers and van Essen 2006], multi-modal input [Dearman et al. 2010], rotation and flexion gestures [Scott et al. 2010], kick gestures [Han et al. 2011], large floor or wall displays [Augsten et al. 2010; Jota et al. 2014], zoom and pan maps [Schöning et al. 2009; Takeuchi 2010], or to rotate a canvas [Sangsuriyachot et al. 2011]. Foot input has, also, found applications in physiotherapy [Paradiso et al. 2004], game interaction [Paelke et al. 2004], interactive dancing [Paradiso and Hu 1997], and mobile device control [Alexander et al. 2012; Crossan et al. 2010].

These and other studies explored the combination of upper and lower limb gestures for computer interaction, unveiling the advantages and drawbacks of feet input. For example, when operating an input device for cursor control, a foot joystick [Pearson and Weiser 1988] and foot controlled trackball [Pakkanen and Raisamo

2004] were slower and more error prone than hand-controlled devices (*e.g.*, mouse, desktop trackball). Besides, seated pedal studies confirmed that foot motion follows Fitts' law but is slower than comparable arm movements [Hoffmann 1991; Pearson and Weiser 1988]. Regarding spatial tasks, Pakkanen and Raisamo [Pakkanen and Raisamo 2004] studied the suitability of foot input for non-accurate spatial operations such as scrolling, moving or resizing objects, and then concluded that feet input is suitable for low accuracy and quick tasks. In accordance with [Pakkanen and Raisamo 2004], Rovers et al. [Rovers and van Essen 2006] shown that feet can be used for relatively simple tasks in hands-busy situations, while Dearman et al. [Dearman et al. 2010] found that simple tap gestures on a foot-pedal were a fast means to select on-screen content. Gobel et al. [Göbel et al. 2013] even designed foot input devices and combined gaze input to support manual interactions in a computer desktop setup, concluding that pan and zoom tasks were performed in a non-fatiguing way using feet. These studies provide encouraging guidelines to the interaction design we wish to adopt, however they all relate to the seated posture or depend a physical foot input device.

Foot-based gestures allow people to control devices comfortably. However, many gestures are too expansive, and thus potentially annoying to other people around the user (*e.g.*, kicking, front toe tap, side heel tap). More importantly, these gestures require users to exercise unbalanced poses, momentarily stand on only a single foot, which can be detrimental to precise 3D manipulation, because of unwanted balance compensation body movements. More subtle gestures, namely single foot heel rotations [Scott et al. 2010] and sequential foot tapping [Augsten et al. 2010; Crossan et al. 2010], have been used to control 2D content and for web browsing as they allow the person to stand in a stable posture. Compared to traditional foot input devices, foot tapping offers slightly diminished efficiency and accuracy for longer selections, yet there is a benefit for short selections that require only four or less foot taps [Crossan et al. 2010]. These studies are complementary to ours since they indicate that heel rotation can be combined with foot tapping to produce useful gestures to interact with digital content, while standing. However, none of these studies focus on contactless 3D manipulation that leverage single foot input to support hand gestures.

### 2.2 3D manipulation with hand and feet gestures

To improve 3D manipulation task performance, many researchers have explored feet as a medium to support hand gestures for selecting modes and controlling a camera in a 3D modelling application [Balakrishnan et al. 1999], selecting menu options using heel rotation and front-foot lift gestures [Zhong et al. 2011], and three-dimensional navigation, selection, manipulation and system control tasks using a depth sensing camera [Simeone et al. 2014]. Regardless of the positive results, the interaction design of these studies revolves around the seated position and/or use foot devices.

More closely related to our work is that of applying foot gestures for 3D manipulation and navigation tasks while standing. The seminal work of [Choi and Ricci 1997] revealed an early application in which feet movements are applied to support three-dimensional manipulation, more specifically, walking and leaning actions to rotate a cylinder in all three-dimensions. On a tabletop environment, Sangsuriyachot et al. [Sangsuriyachot et al. 2011] rotate a cube with subtle foot gestures tracked by a sensor platform where the user stood, while Schöning et al. [Schöning et al. 2009] combined multi-touch hand gestures and foot interaction to assist geospatial operations, as foot input was converted into panning, rotating, tilting and zooming information. These studies are closer to our envisioned hands and feet contactless interaction techniques, yet Jalaliniya et al. [Jalaliniya et al. 2013] work is probably the only study that shares similar interaction contexts for large-scale displays settings, although their system depends on hand wearable and dedicated hardware for foot gesture recognition. Furthermore, their main content are 2D (medical) images and feet are not used to apply any translation or rotation [Jalaliniya et al. 2013].

## 3 INTERACTION DESIGN

We designed three mid-air object manipulation techniques with contextual and mode switching capabilities, not only to shift between different geometric transformations, but to also cope with the number of degrees of freedom required to manipulate a 3D virtual object. To better understand each interaction technique it is necessary to describe their gesture vocabulary.

*Gesture vocabulary.* Following [Pakkanen and Raisamo 2004; Pearson and Weiser 1986] and the design guidelines of [Saunders and Vogel 2015; Scott et al. 2010] that explored the interaction potential of single foot input, our work draws on the strengths of appropriate topologies of foot gestures [Pearson and Weiser 1986] to support mid-air hand gestures for manipulating 3D virtual objects, without depending on expensive dedicated hardware. All considered gestures are performed in the user's physical space while standing in front of a large display. Both hands act as 3D cursors with three possible gesture states: opened, closed, and lasso (Figure 2).

Whenever hand clutching introduces undesired position/angular displacements or hands are not providing proper 3D precision, the user can choose to rely on the dominant foot to support mode switching and to apply relatively small heel rotations to improve manipulation precision. Since the dominant foot provides more effective control input [Pakkanen and Raisamo 2004], the considered feet gesture vocabulary is designed around discrete actions and on continuous position control, namely, a sequence of double taps and heel rotations performed with the dominant foot, while the non-dominant foot lies flat on the floor remaining static to ensure proper weight balance (Figure 2).

The gesture vocabulary (Figure 2) ensures that all three contactless interfaces account for a balanced and comfortable stance. In addition, the user is not limited to a fixed position as the implemented techniques use an in-place posture within the depth sensor acquisition space to demarcate interaction.

*Foot menu.* To provide visual feedback of the dominant foot status, we designed a foot menu (Figure 3). According to the empirical evidence gathered by [Saunders and Vogel 2015], foot targets with an angular size under 22.5° or radial size under 5 cm induce high error rates. We avoid this "fat foot" problem with a fan menu

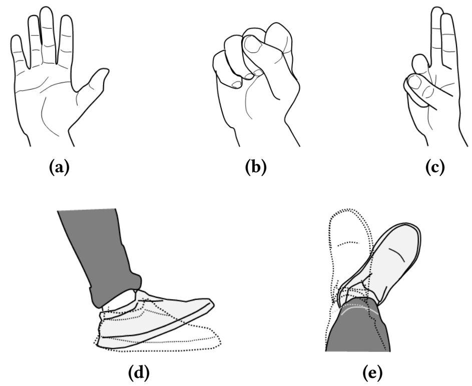


**Figure 2: Gesture vocabulary: (a) open hand showing the palm and all fingers; (b) closed hand forming a fist; (c) lasso with index and middle fingers pointing up; (d) double tapping is a combination of dorsiflexion (movement of the foot up) followed by plantarflexion (movement of the foot down) that is performed twice; (e) heel rotation is either an abduction or adduction, that is the movement of the foot away from (or towards) the body's center line.**

shaped as a semi-circle with two quadrants. Left and right quadrants correspond to the translation and rotation modes, respectively (Figure 3(a)).

Whenever the user wants to perform a manipulation task with the dominant foot, after selecting a transformation, the user must then align the dominant foot closely along the anterior-posterior direction ((Figure 3(b)). The fan menu allows users to select either quadrant with equal foot rotation. When performing a translation or rotation operation, the fan menu also acts like a 1-D slider [Zhong et al. 2011], although the slider interval is placed along the circle's arc length.

## 3.1 Contactless hand gesture interaction (CHE)

The first technique consists of manipulating of 3D virtual objects by only using mid-air hand input. This technique consists on grabbing an object indirectly, moving it to a new position or orientation, and then releasing hand control. Access to all 6 DOF, where translation and rotation is performed simultaneously, is not considered here, as

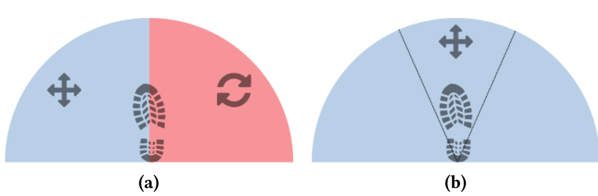


**Figure 3: Foot fan menu. (a) Clutching mode: the left and right quadrants permit to select respectively translation and rotation tasks, while the foot pointer, represented as a boot footprint, indicates which side it is placed; (b) Precision mode: the menu converts to a semi-circle to visually indicate the amount of heel rotation performed by the dominant foot.**

the *CHE* technique only allows objects to be either moved or rotated independently. Within the *CHE* interface, open hand gesture is the default action as it corresponds to moving hand cursors inside the 3D virtual space. No manipulation task is performed while both hands are opened. Mode selection can then be performed by placing the hand cursor above the object and (i) closing the hand to activate translation or (ii) raising the index and middle finger, known as the lasso gesture, to trigger rotation.

Users are able to perform either unconstrained translations by moving the hand in the desired direction, or constrained translations along one of the three main axis contained in the screen plane, as the latter showed promising results in mid-air object manipulation [Mendes et al. 2016]. These axis are selected by placing both hands close to each other in order to form one of the three axis (Figure 4). The same principle applies to constrained rotation. Horizontal movements translate into a rotation around a vertical axis, and vertical movements cause a rotation around a fixed horizontal axis contained in the screen plane.

## 3.2 Clutch-FEETICHE

Our second technique provides the same set of features described previously, although with a significant difference, namely, mode switching between translation or rotation is performed by left or right heel rotation, respectively, which is similar to when driving a car by alternating between pedals with the heel resting on the floor and the foot slightly dorsiflexed. This action is completed after double tapping with the dominant foot. After, selecting the transformation to make, object manipulation is handled by hands like in *CHE*.

## 3.3 Precision-FEETICHE

Our third technique considers the same transformation selection mode than in *Clutch-FEETICHE*. In contrast, object manipulation is handled by the right or left heel rotation. Hands are only used to define the transformation axis while all the visual guidance remains the same. Objects can be moved along or rotated around

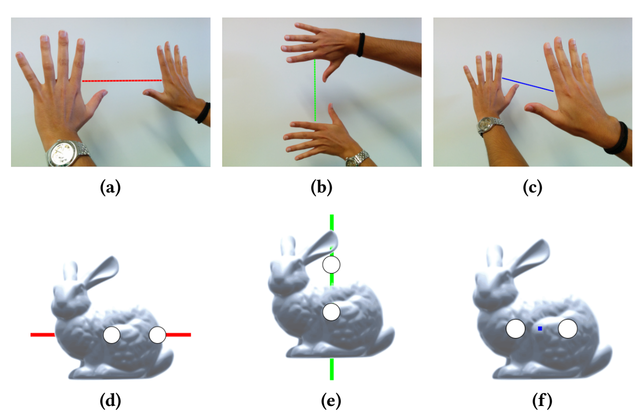


**Figure 4: Hand gestures for axis selection: (a) x-axis; (d) its visual feedback selection; (b) y-axis; (e) its visual feedback selection; (c) z-axis; (f) its visual feedback selection.**

the defined axis by heel rotation. A positive foot rotation, in a clockwise motion, is translated into an positive transformation while a negative transformation is applied if an opposite rotation (counterclockwise) is performed.

# 4 EXPERIMENT 1: ROTATION

We conducted an experiment to compare the performance of *CHE*, *Clutch-FEETICHE* and *Precision-FEETICHE* techniques when rotating a 3D object along an axis.

## 4.1 Participants

Nine unpaid participants (4 female and 5 males) took part in our experiment. Participants' ages varied between 22 and 32 years (mean=25.5, s.d=3.6). All were right-handed. Two participants were regular users of Kinect games.

## 4.2 Apparatus

Our setup comprises a non-invasive and affordable full body tracker using a depth sensing camera (Microsoft Kinect V2), that is placed facing the user (Figure 5). We use the skeleton provided by the Kinect SDK to detect hand positions which are then mapped to the screen. All graphical content was displayed on a 55" LG SUPER UHD TV screen. We developed our prototype using Unity3D engine, with gravity and objects' collision disabled.

Since noisy depth cameras such as the Kinect often fail to deliver either reliable foot tracking or foot tapping input, we use an optical marker placed at in the participant's dominant foot in order to provide a detect height variation to identify foot tapping with respect to an estimated ground level position. The relative position of the marker also provides an estimate of angular heel rotation. The optical marker is attached to the user's shoe either at the toe region or near the foot's instep (Figure 6). The single optical marker also serves to identify a specific dominant foot and not other foot of the same or other user. The marker position is not obstructive as it is always facing the depth sensing camera. While it is arguable that this setup requires special footwear to be worn by surgeons or nurses in operating rooms, surgical procedure already requires specialized attire and thus we feel this requirement is neither too stringent nor too expensive to implement as reflective coating is both cheap and easy to apply.

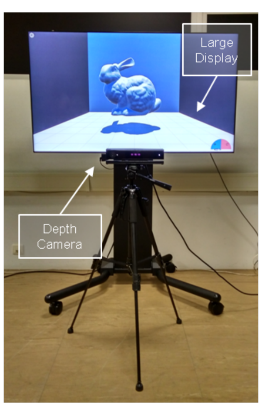


**Figure 5: Hardware setup of the interaction system.**

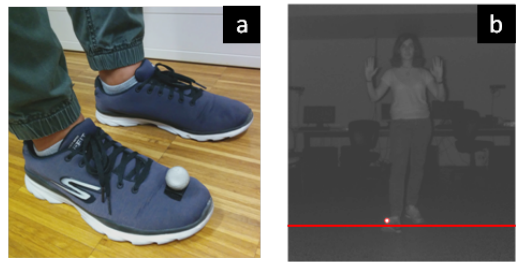


**Figure 6: Motion capture technique to record foot movements. (a) dominant foot with a single optical marker; (b) infrared image with highlighted marker and threshold above which a tap is detected.**

## 4.3 Design

Independent measures are analyzed using a $3 \times 3 \times 2$ repeated measures within-subjects analysis of variance for the factors: *technique* (*CHE*, *Clutch-FEETICHE* and *Precision-FEETICHE*), rotation *axis* (X, Y and Z) and rotation *angle* (*small* (25°) and *large* (100°)).

## 4.4 Task

Participants were asked to put the Stanford Bunny in a required orientation as quickly and accurately as possible. The bunny was the only object in our virtual environment that could be geometrically transformed. The Stanford Bunny model had a semi-transparent replica showing the asked orientation. For each trial, the participant had a total of three attempts, each trial could last up to one minute. A trial was considered successfully completed if the participant placed the object within five degrees of the desired orientation without exceeding the time limit. A trial was deemed as failed if the time limit was exceeded or the wrong transformation was selected by the user (*i.e.*, selecting a translation inside of a rotation task). The participant would move onto the next trial if the trial was successfully completed or if the number of attempts for that trial was exceeded. To help participants during evaluation tasks, the object turned green when the target thresholds were met. Each trial starts with the bunny model at the center of the scene and aligned to the reference system. We then show to the participant the target orientation (Figure 7). Please, also, note that although the task required only one kind of transformation (translation or rotation), none was restricted, as we did not intend to modify any technique in order to accommodate a specific task.

**Figure 7: An example of the rotation task interface where the current and target bunny orientations are shown (Y axis with the large rotation angle).**

### 4.5 Procedure

After answering a demographic questionnaire, we explained the task along with the additional requirements for each interaction technique. The participant then began the experiment. Each technique was evaluated separately and presented in a random order. For each technique, we presented the three rotation *axes* × two *angles* to the participants in random order – there were $3 \times 3 \times 2 = 18$ trials per participant. Subjects could practice three minutes at the start of each technique.

After exercising each interaction technique, participants answered 6-point Likert-scale questions (strongly disagree to strongly agree): (i) *is the task useful?*; (ii) *is the way to perform the task adequate?*; (iii) *was the task easy to perform?*; (iv) *was the task easy to remember?*; and (v) *how successful were you in performing the task?* The average duration of the experiment was 30 minutes.

## 5 RESULTS

We chose as dependent measures the *success rate* (rotation tasks that were not made on first attempt were marked as errors; in doing so, the *success rate* is the percentage of rotation tasks successfully completed on first attempt), *number of failed attempts* (the average unsuccessful attempts made by a participant to select the target (3 at most)), *clutching time* (clutching time is measured from the beginning of the trial, until the expected transformation is successfully selected), and *manipulation time* (manipulation time is measured from the transformation being selected to bunny successfully rotated). We have also analyzed subjective responses. All statistical analyses were multi-way ANOVA. We performed Tukey tests post-hoc when significant effects were found. In the following, we report the results found for each dependent variable in Experiment 1.

### 5.1 Success rate and failed attempts

*Technique* showed a significant main effect ($F_{2,16.26}$=9.96, $p < 0.001$) with a significant *technique* × *axis* interaction ($F_{4,31.63}$=4.25, $p < 0.01$) on *success rate*. In effect, post-hoc tests revealed that when rotating a 3D object around the Z *axis*, *CHE* (mean=23.52, sd=20.78) performed worse in comparison to either *Clutch-FEETICHE* (mean= 83.33, sd=17.71) or *Precision-FEETICHE* (mean=88.23, sd=15.78) techniques ($p < 0.05$). There was no significant *technique* × *axis* × *angle* interaction ($p$=0.39), suggesting that when rotating a 3D object around the *Z* axis, the benefits of either *FEETICHE* techniques over *CHE* are consistent across the different rotation *angles* tested.

Similarly, for SUCCESS RATE, we found significant main effect of *technique* ($F_{2,11.75}$=10.65, $p$=0.002) with a significant *technique* × *axis* interaction ($F_{4,29.71}$=4.36, $p < 0.005$) on *failed attempts* where *CHE* resulted in significantly more failed attempts than either *FEETICHE* technique when rotating the object around the Z axis ($p$<0.05). We also found no significant *technique* × *axis* × *angle* interaction ($p$=.42) suggesting that the benefits of either *FEETICHE* technique over *CHE* are consistent across different rotation *angles* and axes.

### 5.2 Clutching time

*Technique* had a significant main effect ($F_{2,14.37}$=7.82, $p$=0.005) on *clutching time*. Indeed, post-hoc tests revealed that *CHE* was significantly faster than either *Clutch-FEETICHE* or *Precision-FEETICHE* by respectively 42% and 47% ($p < 0.05$). Interestingly, we found that there was no significant interaction either on *technique* × *angle* ($p$=0.88), or *technique* × *axis* ($p$=0.91) or *technique* × *angle* × *axis* ($p$=0.75), suggesting that the benefits of *CHE* are consistent across the different *axes* and *angles*.

### 5.3 Manipulation time

There were significant main effects of *technique* ($F_{2,15.32}$=31.8, $p$<.001) and *axis* ($F_{2,14.51}$=8.68, $p$=0.003) with a significant *technique* × *angle* interaction ($F_{4,17.02}$=7.34, $p$=0.005) on *manipulation time*. Post-hoc tests revealed that *Precision-FEETICHE* was significantly slower than *CHE* for the three rotation *axis* ($p < 0.05$). We also found that the *manipulation time* was significantly higher for *Precision-FEETICHE* than for *Clutch-FEETICHE* when rotating the bunny through the *X axis* ($p < 0.05$). Interestingly, we found that there was no significant *technique* × *angle* × *axis* ($p$=0.73) interaction, suggesting that *Precision-FEETICHE* is consistently slower across the different *axes*.

### 5.4 Subjective results and observations

We recall that participants were asked to rate each technique condition. This is summarized in Table 1. Friedman tests revealed a significant effect technique condition on both suitability and fatigue. Post-hoc comparison using Bonferroni correction showed that *Precision-FEETICHE* is significantly less appropriate than either *CHE* or *Clutch-FEETICHE* techniques ($p < 0.05$). We also found *CHE* as significantly more tiring than either *FEETICHE* techniques ($p < 0.05$).

We correlate these findings with participants' comments that found the *CHE* technique very tiring and requiring more attempts to successfully perform clutching. Some quotes were: *"the hand gestures to release the object were not comfortable after a long usage"*, *"this technique is really tiring!"*. Two participants mentioned that hand gestures were often not recognized and found it to be more difficult to perform constrained rotation when using only hand gestures. One participant even complained that *"the Kinect did not recognize very well my hand gestures and that is why it was difficult to perform the operations."*. Moreover, three participants felt that *Precision-FEETICHE* was uncomfortable as it required more successive double tap gestures as compared to *Clutch-FEETICHE*.

**Table 1: Mean and s.d questionnaire responses, with 1=strongly disagree, and 5 = strongly agree. Friedman tests reported at $p < 0.01$ ($\star$) significance levels. The significant tests are highlighted**

| | **CHE** | | **Clutch-FEETICHE** | | **Precision-FEETICHE** | | **Friedman** |
|---|---|---|---|---|---|---|---|
| | mean | s.d | mean | s.d | mean | s.d | $\chi^2$ |
| Useful | 4.77 | .96 | 4.44 | 1.08 | 4.88 | .68 | .32 |
| Suitability | 4.77 | .96 | 4 | 1.13 | 2.88 | .68 | **9.25**$^\star$ |
| Easiness | 4.11 | 1.15 | 3.55 | 1.08 | 2.66 | .86 | 1.4 |
| Memorability | 4.44 | 1.13 | 4.44 | 1.31 | 4.88 | .68 | 0.75 |
| Performance | 4.33 | .86 | 3.44 | .87 | 2.88 | .82 | 5.42 |
| Fatigue | 5.44 | .87 | 3.88 | 1.19 | 3 | .73 | **11.45**$^\star$ |

In contrast, participants found that *Clutch-FEETICHE* offered a good compromise between hand and feet control, although three participants found double tapping uncomfortable when combined with large heel rotation angles.

## 5.5 Discussion

Our key finding is that *Clutch-FEETICHE* improved *accuracy*, reduced the *number of failed attempts* and decreased the physical effort over conventional *CHE* without compromising *manipulation time*. The performance benefits were consistent across different rotation *angles*. Our analysis suggests also that *Clutch-FEETICHE* is better combined with Z axis manipulations without decreasing performance when applied to the X or Y axes. However, unfortunately *Clutch-FEETICHE* increases the clutching time compared to *CHE*.

It should be noted, that this is not necessarily a demerit, since there is a clear precision/time trade off. Where precision displacement is a clear requirement, even though each *CHE* interaction seemingly takes less time, there will be many repetitions required to compensate for imprecise results. Of course *CHE* is clearly a winner whenever coarse interactions are satisfactory, which arguably is not the case in our scenarios of interest.

Additionally, our findings indicate that *Precision-FEETICHE* is not the most appropriate technique. For instance, even though it improved *accuracy* and reduced the *number of failed attempts* in comparison to *CHE*, it consistently required more time (both for clutching and manipulation) when compared to either *CHE* or *Clutch-FEETICHE* techniques.

Another important result worth highlighting is that FEETICHE effectively solves inadvertent activations and mode switchings, which on and by itself makes interactions more reliable, less frustrating and more comfortable as corroborated by test subjects. This is an important result well sustained by years of multimodal interaction research [Oviatt 1999], that justifies our design choices. Indeed, feet are very effective for clutching (same as in automobile driving tasks) as and enable for greater precision, at the cost of extra task completion times, as they release hands for precise manipulations and argument-setting. A second interesting finding is that it is possible to perform precision tasks using the feet, contrary to established belief and previous result findings, even in constrained settings.

# 6 EXPERIMENT 2: TRANSLATION

In a second experiment we compared *CHE*, *Clutch-FEETICHE* and *Precision-FEETICHE* interaction techniques when translating a 3D object along an axis. To this end, we used the same apparatus as in Experiment 1.

## 6.1 Participants

Nine new unpaid participants (three female + six male) took part in our experiment. Participants' ages varied between 22 and 29 years (mean= 26.1, sd= 4.2). All were right-handed. Two were regular users of Kinect-based console games.

## 6.2 Design

We studied independent variables using $3 \times 2 \times 2$ repeated measures within-subjects analysis of variance for the factors: *technique* (*hands*, *Clutch-FEETICHE* and *Precision-FEETICHE*), translation *axis* (X and Y axis) and translation *distance* (*short* (25 cm) and *long* (100 cm)).

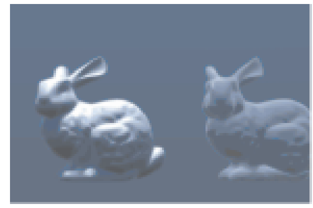

**Figure 8: An example of the translation task interface where the current and target bunny position are shown (X axis with the long distance).**

## 6.3 Task

The tasks required participants to translate a Stanford Bunny model into predefined position as quickly and accurately as possible. The same graphical user interfaces used in Experiment 1 is also considered here. For each trial, a total of three attempts, each lasting no more than one minute, was arranged to every participant. A trial was considered successfully completed if the participant placed the object within two and half centimeters of the desired position without exceeding the time limit. A trial was deemed as failed if, the time limit was exceeded or the wrong transformation was selected by the user, *i.e.*, selecting a rotation inside of a translation task. The participant would start the next trial if it was successfully completed or if the number of attempts was overreached. Each trial starts with the bunny model at the center of the scene and aligned to the reference system. We then show to the participant the target position (Figure 8).

## 6.4 Procedure

After answering a demographic questionnaire, the experiment task was explained for each technique. The participant then began the experiment. Each *technique* is evaluated separately and presented in a random order. Inside each *technique*, the two translation *axis* $\times$ two *distances* were presented in a random order to the participants – overall we have $3 \times 2 \times 2 = 12$ trials per participant. Participants could practice up to three minutes at the beginning of each interface. After exercising each interaction technique, participants responded to the same 6-point Likert-scale questions used in Experiment 1. The average duration of the experiment was 30 minutes.

# 7 RESULTS

In what follows, we used the same dependent measures as in Experiment 1.

## 7.1 Success rate and number of failed attempts

There were significant main effects of *technique* ($F_{2,13.81}$=9.28, $p$=.002), *axis* ($F_{1,5.75}$ = 8.34, $p$=.029) and *distance* ($F_{1,8.04}$=9.06, $p$=.016) on *success rate*. Post-hoc tests showed that *Clutch-FEETICHE* (mean=97.22, s.d=5.44) was significantly more accurate than both *CHE* (mean=82.35, s.d=13) and *Precision-FEETICHE* (mean=67.64, s.d=15.96) ($p$<.05). Interestingly, we found that there was no significant *technique* × *axis* ($p$=.76), nor *technique* × *distance* ($p$=.07) or *technique* × *axis* × *distance* ($p$=.78) interaction, suggesting that the benefits of *Clutch-FEETICHE* are consistent across the different translation *axis* and *distances*.

There were significant main effects of *technique* ($F_{2,15.09}$=11.0428, $p$=.001) and *distance* ($F_{1,8.28}$=9.184, $p$=.015) with *technique* × *distance* ($F_{2,14.33}$=4.38, $p$=.032)) interaction on *number of failed attempts*. Post-hoc tests revealed that, the number of failed attempts was significantly higher with *Precision-FEETICHE* than with *Clutch-FEETICHE* when translating the bunny through the *long* distance (p<.05). We also found, that for *Precision-FEETICHE*; the *number of failed attempts* was significantly higher when translating the 3D object through the *long* distance than trough the *short* distance ($p$<.05). Interestingly, we found that there was no significant *technique* × *axis* × *distance* ($p$=.76) interaction suggesting that the disadvantages of *Precision-FEETICHE* are consistent across different *axes*.

## 7.2 Clutching time

There were significant main effects of *technique* ($F_{2,14.23}$= 5.30, $p$= .019) on *clutching time*. Post-hoc tests revealed that *CHE* was significantly faster than both *Clutch-FEETICHE* and *Precision-FEETICHE* by respectively 54.31% and 56.40% ($p$<.05). We also found that there were no significant *technique* × *axis* ($p$=.44) nor *technique* × *distance* ($p$=.49) or *technique* × *axis* × *distance* ($p$=.19) interactions suggesting that the benefits of *CHE* are consistent across different *axis* and *distances*.

## 7.3 Manipulation time

We observed significant main effects of *technique* ($F_{2,15.59}$=30.09, $p$<.0001), *distance* ($F_{1,8.36}$=40.04, $p$=.0002) with *axis* × *distance* interaction ($F_{1,7.86}$=11.96, $p$=.008) on *manipulation time*. Post-hoc tests revealed that *CHE* (respectively *Clutch-FEETICHE*) was significantly faster than both *FEETICHE* techniques (respectively *Precision-FEETICHE*). We also found no significant *technique* × *axis* ($p$=.58) nor *technique* × *distance* ($p$=.07) or *technique* × *axis* × *distance* ($p$=.36) interactions suggesting that the benefits of *CHE* and *Clutch-FEETICHE* were consistent across different *axes* and *distances*.

## 7.4 Subjective results and observations

We recall that participants were asked to rate each condition. These results are summarized in Table 2. Friedman tests revealed a significant effect of the technique on suitability, ease of use, performance and fatigue. Post-hoc comparison using Bonferroni correction showed that *Precision-FEETICHE* induced significantly a decrease in the perceived suitability and easiness comparing to *CHE* ($p$ < 0.05). In contrast, *CHE* induces significantly an increase in the fatigue comparing to both *FEETICHE* techniques ($p$ < 0.05).

## 7.5 Discussion

Our key finding is that *Clutch-FEETICHE* improved the *accuracy* and decreases the physical effort over conventional *CHE* technique. However, it amplified consistently the time (clutching time and manipulation time). Our findings also show that *Precision-FEETICHE* is not at all appropriate. For instance, due to its increased time requirements, *Precision-FEETICHE* induced higher perceived difficulty in

**Table 2: Mean and s.d questionnaire responses, with 1=strongly disagree, and 5 = strongly agree. Friedman tests are reported at $p$<.05 (★) significance levels. The significant tests are highlighted.**

| | CHE | | Clutch-FEETICHE | | Precision-FEETICHE | | Friedman |
|---|---|---|---|---|---|---|---|
| | mean | s.d | mean | s.d | mean | s.d | $\chi^2$ |
| Useful | 5.11 | .51 | 4.55 | .98 | 4.66 | .56 | 1.85 |
| Suitability | 5.11 | .51 | 4.11 | 1.05 | 3 | .56 | **12.19★** |
| Easiness | 4.55 | 1.22 | 4.55 | 1.08 | 3.11 | .76 | **6.06★** |
| Memorability | 4.77 | 1.12 | 4.55 | 1.22 | 4.77 | 1.02 | 1 |
| Performance | 4.55 | 1.18 | 4.22 | 1.02 | 3 | .56 | **7.8★** |
| Fatigue | 4.77 | 1.12 | 3.66 | 1.22 | 3.11 | .82 | **6.25★** |

completing the translation task. However, again precision comes at a cost and this is expected in the light of the results of the previous experiment. Again, people found performing the task less fatiguing using either variant of FEETICHE than *CHE*. And these findings justify our confidence on the results and vindicate our approach to more precise interactions. While we felt that translations along Z axis are more difficult to accomplish due to human limitations at discerning between different close depths, research reported to ISMAR and VR suggest that auxiliary cues coupled with stereoscopic displays may render such manipulations more effective.

We conclude that the proposed gesture vocabulary for 3D manipulation is based on simple hand and foot movements and users from operating over long periods. Other postures could also benefit from the proposed gesture vocabulary as it can be easily adapted to either seated postures or public displays. Yet, the foot vocabulary is not extendable to walking as tapping and heel rotation are bio-mechanically unfeasible during the gait cycle.

Similarly to [Saunders and Vogel 2016], and since foot input is always available, *FEETICHE* interaction techniques also enable mode shifting actions without disrupting task flow. Note that intense task focus without breaks does not occur in our experimental setting.

The devised motion capture technique is simple yet robust that allows us to track double tapping and heel rotation gestures of the dominant foot. Rather than relying on complicated image and depth map processing techniques [Gritti et al. 2014], our system takes a simple image processing approach to recognize foot gestures, using intensity features extracted from image data.

# 8 CONCLUSIONS AND NEXT STEPS

We presented *FEETICHE*, two interaction techniques that combine hand and foot input for contactless manipulation of 3D virtual objects. *FEETICHE* techniques are compatible with a low-cost depth sensing device in a large-display setting, while aiming to augment hand gestures without being too intrusive yet relies on comfortable foot-based interactions. Through two experiments we have demonstrated the benefits and the limitations of *FEETICHE* techniques for rotation and translation tasks. Our findings demonstrate that *Clutch-FEETICHE* offered a good compromise between hand and feet control, by improving the accuracy and reducing the physical effort over conventional *CHE* technique. In contrast, our findings indicate that *Precision-FEETICHE* is less suitable for contactless 3D object manipulation.

Given the results and participant feedback, our *FEETICHE* approach works for its intended purpose: to effectively support controller-less hand input using comfortable feet gestures to operate a contactless interface for 3D manipulation, while standing. Therefore, *FEETICHE* is a promising approach to a wide range of contactless scenarios as our findings encourage the use of the proposed interaction techniques for controlling real applications. To this end, we plan to build a fully working system after adapting the design above and wish to evaluate the advantages of the interaction techniques that combine hands and feet in a real setting such as surgical navigation, where contactless interfaces are welcomed to maintain a sterile environment. Finally, we hope that our results will further deeper understanding of using feet input for contactless 3D object manipulation which subsequently will be useful to researchers and practitioners for improving user experience in this scenario.

## ACKNOWLEDGMENTS

This work was partially supported by the Portuguese Foundation for Science and Technology (FCT) through grants UID/CEC/50021/2019, IT-MEDEX PTDC/EEI-SII/6038/2014, STREACKER UTAP-EXPL/CA/0065/2017 and SFRH/BD/136212/2018.

## REFERENCES

Jason Alexander, Teng Han, William Judd, Pourang Irani, and Sriram Subramanian. 2012. Putting Your Best Foot Forward: Investigating Real-world Mappings for Foot-based Gestures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12)*. ACM, New York, NY, USA, 1229–1238. https://doi.org/10.1145/2207676.2208575

Ferran Argelaguet and Carlos Andujar. 2013. A survey of 3D object selection techniques for virtual environments. *Computers & Graphics* 37, 3 (2013), 121 – 136. https://doi.org/10.1016/j.cag.2012.12.003

Thomas Augsten, Konstantin Kaefer, René Meusel, Caroline Fetzer, Dorian Kanitz, Thomas Stoff, Torsten Becker, Christian Holz, and Patrick Baudisch. 2010. Multitoe: High-precision Interaction with Back-projected Floors Based on High-resolution Multi-touch Input. In *Proceedings of the 23Nd Annual ACM Symposium on User Interface Software and Technology (UIST '10)*. ACM, New York, NY, USA, 209–218. https://doi.org/10.1145/1866029.1866064

Ravin Balakrishnan, George Fitzmaurice, Gordon Kurtenbach, and Karan Singh. 1999. Exploring Interactive Curve and Surface Manipulation Using a Bend and Twist Sensitive Input Strip. In *Proceedings of the 1999 Symposium on Interactive 3D Graphics (I3D '99)*. ACM, New York, NY, USA, 111–118. https://doi.org/10.1145/300523.300536

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

I. Choi and C. Ricci. 1997. Foot-mounted gesture detection and its application in virtual environments. In *1997 IEEE International Conference on Systems, Man, and Cybernetics. Computational Cybernetics and Simulation*, Vol. 5. 4248–4253 vol.5. https://doi.org/10.1109/ICSMC.1997.637367

Andrew Crossan, Stephen Brewster, and Alexander Ng. 2010. Foot Tapping for Mobile Interaction. In *Proceedings of the 24th BCS Interaction Specialist Group Conference (BCS '10)*. British Computer Society, Swinton, UK, UK, 418–422.

David Dearman, Amy Karlson, Brian Meyers, and Ben Bederson. 2010. Multi-modal Text Entry and Selection on a Mobile Device. In *Proceedings of Graphics Interface 2010 (GI '10)*. Canadian Information Processing Society, Toronto, Ont., Canada, Canada, 19–26.

Fabian Göbel, Konstantin Klamka, Andreas Siegel, Stefan Vogt, Sophie Stellmach, and Raimund Dachselt. 2013. Gaze-supported Foot Interaction in Zoomable Information Spaces. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems (CHI EA '13)*. ACM, New York, NY, USA, 3059–3062. https://doi.org/10.1145/2468356.2479610

A. P. Gritti, O. Tarabini, J. Guzzi, G. A. Di Caro, V. Caglioti, L. M. Gambardella, and A. Giusti. 2014. Kinect-based people detection and tracking from small-footprint ground robots. In *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*. 4096–4103. https://doi.org/10.1109/IROS.2014.6943139

Teng Han, Jason Alexander, Abhijit Karnik, Pourang Irani, and Sriram Subramanian. 2011. Kick: Investigating the Use of Kick Gestures for Mobile Interactions. In *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services (MobileHCI '11)*. ACM, New York, NY, USA, 29–32. https://doi.org/10.1145/2037373.2037379

Ken Hinckley, Randy Pausch, John C. Goble, and Neal F. Kassell. 1994. A Survey of Design Issues in Spatial Input. In *Proceedings of the 7th Annual ACM Symposium on User Interface Software and Technology (UIST '94)*. ACM, New York, NY, USA, 213–222. https://doi.org/10.1145/192426.192501

Errol R. Hoffmann. 1991. A comparison of hand and foot movement times. *Ergonomics* 34, 4 (1991), 397–406. https://doi.org/10.1080/00140139108967324

Christian Holz and Andrew Wilson. 2011. Data miming: inferring spatial object descriptions from human gesture. In *CHI 2001*.

Shahram Jalaliniya, Jeremiah Smith, Miguel Sousa, Lars Büthe, and Thomas Pederson. 2013. Touch-less Interaction with Medical Images Using Hand & Foot Gestures. In *Proceedings of the 2013 ACM Conference on Pervasive and Ubiquitous Computing Adjunct Publication (UbiComp '13 Adjunct)*. ACM, New York, NY, USA, 1265–1274. https://doi.org/10.1145/2494091.2497332

Ricardo Jota, Pedro Lopes, Daniel Wigdor, and Joaquim Jorge. 2014. Let's Kick It: How to Stop Wasting the Bottom Third of Your Large Screen Display. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 1411–1414. https://doi.org/10.1145/2556288.2557316

Daniel Simões Lopes, Pedro Duarte de Figueiredo Parreira, Soraia Figueiredo Paulo, Vitor Nunes, Paulo Amaral Rego, Manuel Cassiano Neves, Pedro Silva Rodrigues, and Joaquim Armando Jorge. 2017. On the utility of 3D hand cursors to explore medical volume datasets with a touchless interface. *Journal of Biomedical Informatics* 72, Supplement C (2017), 140 – 149. https://doi.org/10.1016/j.jbi.2017.07.009

A. Martinet, G. Casiez, and L. Grisoni. 2010. The design and evaluation of 3D positioning techniques for multi-touch displays. In *2010 IEEE Symposium on 3D User Interfaces (3DUI)*. 115–118. https://doi.org/10.1109/3DUI.2010.5444709

Daniel Mendes, Filipe Relvas, Alfredo Ferreira, and Joaquim Jorge. 2016. The Benefits of DOF Separation in Mid-air 3D Object Manipulation. In *Proceedings of the 22Nd ACM Conference on Virtual Reality Software and Technology (VRST '16)*. ACM, New York, NY, USA, 261–268. https://doi.org/10.1145/2993369.2993396

Sharon Oviatt. 1999. Ten Myths of Multimodal Interaction. *Commun. ACM* 42, 11 (Nov. 1999), 74–81. https://doi.org/10.1145/319382.319398

Volker Paelke, Christian Reimann, and Dirk Stichling. 2004. Foot-based Mobile Interaction with Games. In *Proceedings of the 2004 ACM SIGCHI International Conference on Advances in Computer Entertainment Technology (ACE '04)*. ACM, New York, NY, USA, 321–324. https://doi.org/10.1145/1067343.1067390

Toni Pakkanen and Roope Raisamo. 2004. Appropriateness of Foot Interaction for Non-accurate Spatial Tasks. In *CHI '04 Extended Abstracts on Human Factors in Computing Systems (CHI EA '04)*. ACM, New York, NY, USA, 1123–1126. https://doi.org/10.1145/985921.986004

Joseph A. Paradiso and Eric Hu. 1997. Expressive Footwear for Computer-Augmented Dance Performance. In *Proceedings of the 1st IEEE International Symposium on Wearable Computers (ISWC '97)*. IEEE Computer Society, Washington, DC, USA, 165–.

Joseph A. Paradiso, Stacy J. Morris, Ari Y. Benbasat, and Erik Asmussen. 2004. Interactive Therapy with Instrumented Footwear. In *CHI '04 Extended Abstracts on Human Factors in Computing Systems (CHI EA '04)*. ACM, New York, NY, USA, 1341–1343. https://doi.org/10.1145/985921.986059

G. Pearson and M. Weiser. 1986. Of Moles and Men: The Design of Foot Controls for Workstations. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '86)*. ACM, New York, NY, USA, 333–339. https://doi.org/10.1145/22627.22392

G. Pearson and M. Weiser. 1988. Exploratory Evaluation of a Planar Foot-operated Cursor-positioning Device. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '88)*. ACM, New York, NY, USA, 13–18. https://doi.org/10.1145/57167.57169

A.F. Rovers and H.A. van Essen. 2006. Guidelines for haptic interpersonal communication applications: an exploration of foot interaction styles. *Virtual Reality* 9, 2 (01 Mar 2006), 177–191. https://doi.org/10.1007/s10055-005-0016-0

Nuttapol Sangsuriyachot, Haipeng Mi, and Masanori Sugimoto. 2011. Novel Interaction Techniques by Combining Hand and Foot Gestures on Tabletop Environments. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces (ITS '11)*. ACM, New York, NY, USA, 268–269. https://doi.org/10.1145/2076354.2076409

William Saunders and Daniel Vogel. 2015. The performance of indirect foot pointing using discrete taps and kicks while standing. In *Proceedings of Graphics Interface 2015 (GI 2015)*. Canadian Human-Computer Communications Society, Toronto, Ontario, Canada, 265–272. https://doi.org/10.20380/GI2015.34

William Saunders and Daniel Vogel. 2016. Tap-Kick-Click: Foot Interaction for a Standing Desk. In *Proceedings of the 2016 ACM Conference on Designing Interactive Systems (DIS '16)*. ACM, 323–333. https://doi.org/10.1145/2901790.2901815

Johannes Schöning, Florian Daiber, Antonio Krüger, and Michael Rohs. 2009. Using Hands and Feet to Navigate and Manipulate Spatial Data. In *CHI '09 Extended Abstracts on Human Factors in Computing Systems (CHI EA '09)*. ACM, New York, NY, USA, 4663–4668. https://doi.org/10.1145/1520340.1520717

Jeremy Scott, David Dearman, Koji Yatani, and Khai N. Truong. 2010. Sensing Foot Gestures from the Pocket. In *Proceedings of the 23Nd Annual ACM Symposium on User Interface Software and Technology (UIST '10)*. ACM, New York, NY, USA, 199–208. https://doi.org/10.1145/1866029.1866063

A. L. Simeone, E. Velloso, J. Alexander, and H. Gellersen. 2014. Feet movement in desktop 3D interaction. In *2014 IEEE Symposium on 3D User Interfaces (3DUI)*. 71–74. https://doi.org/10.1109/3DUI.2014.6798845

Yuichiro Takeuchi. 2010. Gilded Gait: Reshaping the Urban Experience with Augmented Footsteps. In *Proceedings of the 23Nd Annual ACM Symposium on User Interface Software and Technology (UIST '10)*. ACM, New York, NY, USA, 185–188. https://doi.org/10.1145/1866029.1866061

Eduardo Velloso, Dominik Schmidt, Jason Alexander, Hans Gellersen, and Andreas Bulling. 2015. The Feet in Human–Computer Interaction: A Survey of Foot-Based Interaction. *ACM Comput. Surv.* 48, 2, Article 21 (Sept. 2015), 35 pages. https://doi.org/10.1145/2816455

Vinayak, Sundar Murugappan, HaiRong Liu, and Karthik Ramani. 2013. Shape-It-Up: Hand gesture based creative expression of 3D shapes using intelligent generalized cylinders. *Computer-Aided Design* 45, 2 (2013), 277 – 287. https://doi.org/10.1016/j.cad.2012.10.011 Solid and Physical Modeling 2012.

K. Zhong, F. Tian, and H. Wang. 2011. Foot Menu: Using Heel Rotation Information for Menu Selection. In *2011 15th Annual International Symposium on Wearable Computers*. 115–116. https://doi.org/10.1109/ISWC.2011.10