Freehand Grasping: An Analysis of Grasping for Docking Tasks in Virtual Reality

Figure 1: Virtual environment for grasp patterns elicitation study. *Equant, Prolate, Oblate, Bladed* virtual objects were randomly positioned on the table, each of the objects having a corresponding target object (highlighted in green). At the end of the docking task all objects had to be in the target position and rotation.

ABSTRACT

Natural and intuitive interaction in VR as grasping virtual objects, is still a significant challenge and while recent studies have begun to explore interactions that aim to seamlessly create virtual environments that mimic reality as closely as possible, the dexterous versatility of the human grasp poses significant challenges for usable and intuitive interactions. At present the design considerations for creating natural grasping based interactions in VR are usually drawn from the body of historical knowledge presented for real object grasping. While this may be suitable for some applications, recent work has shown that users in VR grasp virtual objects differently than they would grasp real objects. Therefore, these interaction assumptions may not be directly applicable in furthering the natural interface for users of VR, presenting an absence of knowledge on how users intuitively grasp virtual objects. To begin to address this, we present two experiments where participants (N=39) grasped 16 virtual objects categorised by shape in a mixed docking task exploring rotation, placement and target location. We report on a *Wizard of Oz* methodology and extract grasp types, grasp category and grasp dimension. We further provide insights into virtual object categorisation for assessing interaction patterns and how these could be used for developing natural and intuitive grasp models by parameterizing grasp types found in these experiments. Our results are of value to be taken forward into a framework of recommendations for grasping interactions and thus begin to bridge the gap in understanding natural grasping patters for VR object interactions.

Keywords: Hand Interaction, Grasping Virtual Objects, Virtual Reality Interaction

1 INTRODUCTION

Virtual Reality (VR) headsets are currently experiencing an unprecedented growth in consumer availability [56], changing the way people communicate [58], play [46], work [47] and learn [40] via achieving high levels of user immersion [6].

These applications aim to achieve high user performance, and therefore need to mimic reality as closely as possible and enable highly immersive environments and natural interactions [36]. To achieve this, researchers have focused on developing VR interactions that use the human hand as an input tool [1], taking advantage of its multiple degrees of freedom [63] and the human ability to use hands for acquiring and manipulating objects with ease [57].

Hand interactions relying on predefined gestures have been extensively explored during the past years [25], however, researchers showed that these interactions are predominantly designed for optimal recognition rather than for naturalness, often not being intuitive enough [59]. Considering this, researchers focused on interactions that act physically on the virtual objects [68], also known as grasping.

Grasping is the primary form of manual interaction between humans and the physical world [2] and is defined as every static posture at which an object can be held securely with a single hand [33]. While grasping real objects has been highly explored before [61], grasping virtual objects in a natural manner is still a challenge [60].

Tasks have been explored in the context of real object grasping, showing that grasping patterns are primarily influenced by object characteristics [32] and the task to be performed [31]. Moreover, it has been shown that the task to be performed also plays an important role for virtual hand interactions, researchers designing a variety of tasks such as pointing, reaching and tilting [53] or moving, rotating or scaling [3] were used to assess user performance and patterns for interaction and showed that different tasks require different interaction approaches when manipulating virtual objects [7,59]. However, the influence of task properties on freehand grasping interaction patterns has not been yet defined. Furthermore, a rich understanding of virtual object categorisation inspired by Zingg's methodology [76] is also novel for this work.

To define intuitive gestural preferences, researchers conducted elicitation studies [28] to get insights into how to achieve interactions that are natural and intuitive to the users [74]. Following this approach, and aiming to begin to address the gap previously men-

tioned, we conduct two elicitation experiments and involve the user in defining grasp patterns for virtual object manipulation docking tasks in VR. We ask users to intuitively grasp virtual objects and perform mixed docking tasks (placement and orientation).

We further discuss our results which provide insights for designers of interactive 3D hand interactions for determining the most suitable grasp type that should be used for triggering interaction based on the virtual object characteristics and the task to be performed. We contribute by providing insights into virtual object categorisation using Zingg's methodology [76] for assessing interaction parameters as well as the influence of object characteristics on grasping patterns and how these could be further developed into natural and intuitive grasp models.

This paper is structured as follows: Section 2 presents the background and related work, notably literature on interaction in VR, grasping virtual objects and elicitation studies. Section 3 presents the methodology for the first experiment and Section 4 presents the methodology for the second experiment. Section 5 presents metrics, namely labelling the grasps methodology along with the grasping metrics. Section 6 presents the results in terms of grasp types and grasp dimension, Section 7 presents discussion around the results and design recommendations and conclude in Section 8.

2 BACKGROUND

2.1 Interaction in Virtual Reality

With VR being highly used for training [54] or simulations [67], creating highly immersive environments with high visual and interaction fidelity is now of huge interest for researchers [36]. Interactions in VR were initially focused around using physical props [37], however, more recent work focused around achieving high fidelity interactions, that are known to play an important role in increasing immersion in virtual environments [39].

Recent VR systems use the hand, the most powerful tool through which we interact with the surrounding world, as the main interaction tool [1]. However, when creating bare hand interactive systems, aiming to provide seamless and intuitive interactions, researchers have focused on a natural approach known as grasping [2].

This natural interaction with objects in VR plays an important role in increasing immersion in virtual environments [39], however is still one of the ongoing interaction challenges in VR [60]. The next section introduces studies that used docking tasks to assess interaction in virtual environments.

2.1.1 Docking Tasks

Target acquisition is one of the most elementary interactions in 3D environments [29]. However, the interaction choice, a critical factor in usability and performance of VR systems, is highly influenced by the task to be performed after target acquisition stage [52]. Therefore, researchers have used variations of tasks such as placement, orientation and docking to assess aspects of virtual object manipulations [30, 52]. Docking is a more sophisticated task that combines point or volume matching with orientation matching [64] and has been highly used by researchers for assessing interaction metrics under different experiment conditions in VR environments. Vuibert *et al.* [71] used a docking task to assess gross motion and then fine-tuning once near the target and compare different interaction methods based on accuracy and completion time. Bai et al. [4] implemented rotation and translation tasks for mid-air object 3D manipulation and proposed a comprehensive repertoire of 3D manipulation operations to manipulate objects. Boritz *et al.* [9] created placement and orientation tasks to assess different visual feedback modes and showed that target position and orientation on a docking task have a significant effect upon user performance in VR.

While we show that docking tasks have previously been used in VR interaction studies, grasping interaction patterns while perform-

ing docking tasks were not explored before. Next section presents grasping literature.

2.2 Grasping Virtual Objects

Researchers have investigated human's approach to grasping real objects, aiming at understanding certain aspects of human hand usage [61], as well as classifying grasps in a discrete set of types [23, 34]. However, this knowledge cannot be directly applied to grasping virtual objects, as it has been identified that users interact with virtual objects differently than they interact with real objects [51]. Therefore, virtual grasping has been explored as a technical challenge [10, 11], with the aim of enabling users to manipulate virtual objects as if they were real [2, 13]. However, current approaches predominantly rely on computed grasps, not taking into account humans' grasping actions [41]. However, these limitations can be addressed by closely analysing human actions while grasping objects in VR [69]. Accordingly, Blaga *et al.* [8] analysed human actions while grasping virtual objects by investigating the influence of thermal visual cues on grasping in VR, to understand how virtual environments influence grasping approaches. However, this study only investigated one virtual object (the mug) and was only looking at thermal visual cues. To understand how other interaction metrics that are known to influence grasping in real object manipulation, such as object characteristics [32] and task properties [31], we conduct two elicitation experiments. The next section introduces elicitation studies and how they play a key role in defining interaction patterns.

2.3 Elicitation Studies

Gesture elicitation methodology [73, 75] has proven immensely resourceful for designers to form an understanding of users' mental models of gesture interaction. Since its first implementation for multi-touch gestures [75], the methodology has been reapplied for a variety of gesture types and applications [18], focusing on involving the user and revealing their preferences for interactive gestures, therefore accumulating important interaction design knowledge. Wobbrock *et al.* followed this methodology to define gestures for tabletop interactions, conducting an elicitation study that assesses users' gesture when executing specific tasks. Before the study began, the three authors independently designed their own gestures, and proved that, even combined, authors only covered 60.9% of the set proposed by the users, showing the impact of elicitation on understating interaction patterns. Elicitation studies have also been used to define intuitive gesture-based interfaces for applications such as: multi-touch gestures for mobile devices [55], communicating between multiple devices, such as mobile-phones, public displays, tablets and tabletops [45] or interactive storytelling in VR scenarios [44].

One of the outcomes of these elicitation studies, is a set of recommendable gestures together with an estimation of users' consensus or agreement [73], playing an essential role and being a first step into improving natural interactions. However, elicitation studies for grasping patterns when performing docking tasks have not been performed yet. We therefore conduct two experiments to understand the influence object characteristics and task properties have on grasp patterns. The next section introduces the methodology for the experiments.

3 EXPERIMENT 1 [E1]

3.1 Apparatus

We built a custom experimental framework using the Oculus DK2 VR headset and the Leap Motion device. Additionally, we used a Logitech Pro 1080p HD webcam positioned on top of the Oculus DK2, based on the methodology of [8] to record participants' hand from the user's perspective at all times. The system was developed using C#, Unity 2018.2 and Leap Motion 4.0 SDK. The Leap Motion
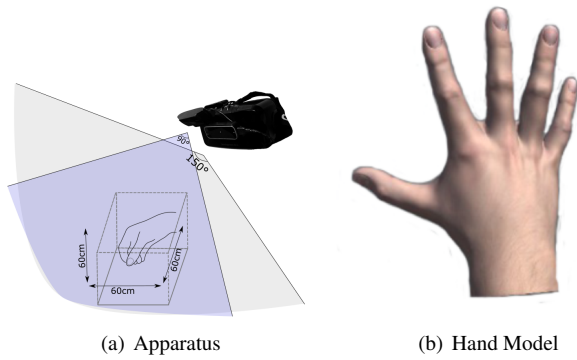
(a) Apparatus   (b) Hand Model

Figure 2: System configuration, with (a) displaying the overall test configuration and (b) displaying the virtual hand used for interaction

Controller was attached to the HMD, facing the user's hands. The virtual interaction space was $60\,cm \times 60\,cm \times 60\,cm$ and the starting position was the same for each participant. The setup and equipment are shown in Figure 2 (a).

3.2 Hand Representation

The use of avatars and avatar representation in VR has received significant attention from the research community exploring how it affects the sense of body ownership and agency [43]. While additional visual aspects such as human-likeness [48], or gender [65] of the virtual representation have shown to influence ownership illusion as well as user performance, when comparing between androgynous human 3D models for grasping virtual objects, no significant differences were found [8]. Therefore, we chose an androgynous model, following Schwind *et al.* recommendation for avoiding noticeable gender characteristics in human hands [65] as shown in Figure 2 (b).

3.3 Referents

Previous user elicitation studies have used pictorial [74] or animated [59] referents to encourage participants to develop their own set of gestures based on showing the effects (*referents*) these will have on the system. When grasping a real objects, humans know how an object should be grasped based on its characteristics and the task to be performed [32]. Therefore, we present known real life object in VR, along with orientation and placement tasks as *referents* in our elicitation experiments, as they inherently provide insights into how the object should be grasped.

3.3.1 Virtual Objects

A representative set of 16 daily life virtual objects from the *"Yale–Carnegie Mellon University–Berkeley Object and Model Set"* [16], a reference for objects of daily life [19] were used. We followed the methodology of [31,32] for investigating grasping patterns when manipulating real objects and categorised our subset of virtual objects in four categories by applying Zingg's methodology [76].

Zingg defined "A" as the longest dimension of an object and "C" the shortest, being "B" the remaining dimension. He defined a constant R to define the relationship between dimensions and categorise the object; determining that the value at which one typically regards two axes to be different is about R=3/2 [76]. Based on these parameters, he defined four shape categories: *Equant*, *Prolate*, *Oblate* and *Bladed*. We applied this methodology and categorised objects from our subsets in these four categories, based on their dimensions in the virtual environment. Zingg's methodology with each category's definition along with the selected objects are presented in Figure 3.

(a) Equant shaped objects

(b) Prolate shaped objects

(c) Oblate shaped objects
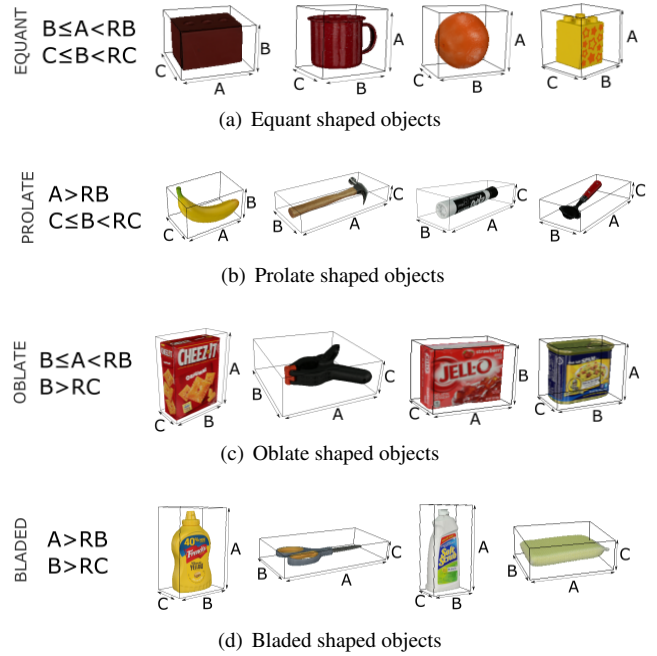
(d) Bladed shaped objects

Figure 3: Selected objects from [16] divided by Zingg's categories and object characteristics [76] : Equant, Prolate, Oblate and Bladed

To our knowledge this is the first example of Zingg's methodology in VR objects.

3.3.2 Task

Task intention has shown to influence grasp patterns when interacting with real objects, presenting different grasp choices for the same object when the task to be performed changes [32]. Moreover, VR systems where natural interactions have an impact on overall performance are task-oriented [50].

When manipulating virtual 3D content, placement, orientation and docking (6 Degrees of Freedom) are fundamental tasks for which the choice of interaction gestures is critical for usability and performance [52]. Therefore, to understand the pattern of grasp choice in VR, we designed a task for 3D object manipulation, which consisted in a docking task (placement and orientation of a virtual object to a target position and rotation) to understand the intuitive grasp choices users take when performing docking tasks in VR.

Virtual Object Rotation: Research shows that there are some limitations to rotating virtual objects compared to real objects, and that a simple task can exceed 20 seconds whereas the same kind of rotation can be accomplished with real objects and with some VR interfaces in less than two seconds [72]. To assess if this has an influence on grasping patterns, we rotated the virtual objects on the horizontal plane (azimuth) by a value of 30°to create a realistic effect in the virtual environment (objects positioned and rotated on Y axis on the virtual table, not floating on the table as in the case of them being rotated on X or Z axis) with objects being ergonomically reachable by the hand.

Target Categories: In Section 3.3.1 we present the criteria for selecting the virtual objects based on their shape. The chosen objects span a variety of manipulation applications and challenges so we further categorised the objects based on their daily use, as in the work of Calli *et al.* [15]. While the objects to be grasped were randomly positioned on the virtual table, the target objects were categorised based on their daily use as follows:

(a) Tools   (b) Groceries   (c) Fruits

Figure 4: Target categories *Tools, Groceries, Fruits* as presented in Section 3.3.2 showing the object targets categorised by their daily usage.

- **Tools:** hammer, scissors and clamp targets were positioned in this target zone (left side shown in Figure 1).

- **Groceries:** crackers, mustard, cleanser, gelatin, spoon, brick, marker, lego, mug and meat can targets were positioned in this target zone (centre of the scene as shown in Figure 1).

- **Fruits:** banana and orange targets were positioned in this target zone (right side as shown in Figure 1).

In Figure 5 we show an example docking task for the *Cracker Box* virtual object. The object needs to be rotated and placed at the target position and rotation highlighted in green colour. This task is applied to each of the 16 virtual objects used in the experiment.

3.3.3 Environment

The experiment was conducted in a controlled environment under laboratory conditions. The test room was lit by a 2700k (*warm white*) fluorescent with no external light source. While VR tasks can be associated with simulator sickness, docking tasks are designed to be completed in a static scene with the user focusing on a single object at a time [64]. Accordingly, the virtual environment in this scene is static, showing a virtual desk with its surface aligned to a seating position, and objects positioned on the virtual table, randomly arranged and at a distance from each other of 10cm in each direction in the virtual space, mimicking reality as closely as possible (see Fig. 1)

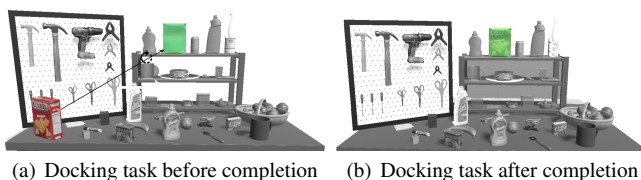(a) Docking task before completion   (b) Docking task after completion

Figure 5: Example of one of docking task for the *Cracker Box* virtual object from *Equant* category. (a) Docking before completion shows the task before the user grasps it and translate + rotate it to the target position (highlighted in green); (b) Docking task after completion shows the task after the target was translated and rotation to the target position (overlaying the green area).

3.4 Participants

A total of 39 participants (25 male and 14 female) from a population of university students and staff members volunteered to take part in this study. Participants ranged in age from 19 to 47 (M=27.69, SD=6.66). All participants were right-handed, to ensure they interacted with the virtual objects under the same conditions.

All participants performed all the experiment tasks. Participants completed a standardised consent form and were not compensated.

Visual acuity of participants was measured using a Snellen chart, each participant was also required to pass an Ishihara test to check for colour blindness. Participants with colour blindness and/or non corrected visual acuity of < 0.80 (where 20/20 is 1.0) were not included in this study.

Participants were asked to self-assess their level of experience with VR systems and gesture recognition systems. Regarding VR systems, 21 reported being novice to the technology, 16 reported having an average level of experience and 2 self-labelled themselves as experts. Regarding gesture-recognition systems, 27 reported being novice to the technology, 10 reported an average level of experience and 2 self-labelled themselves as experts.

3.5 Protocol

A *Wizard of Oz* method is defined as the experiment approach where the interaction is mediated by a human operator to allow the user more freedom of expression or constrain the interaction in a systematic way [24]. These approaches proved useful for gathering data as a basis for theories in human-computer interaction studies [20]. In order to collect grasp data during docking tasks in VR, we need to follow a *Wizard of Oz* approach to trigger interaction with virtual objects for every grasping action performed by the user. This way we allow the user to choose how to grasp the virtual object and the interaction will be triggered for every grasp pose to complete the task.

3.5.1 Pre-test

Prior to the study, participants were given a written informed consent where the test protocol and main aim of the study was described. Additionally, participants completed a pre-test questionnaire enquiring about their background level of experience with VR systems and hand recognition sensors.

3.5.2 Training

Participants underwent initial hand interaction and task training to familiarise themselves with the VR environment and hand interaction space. This training task was a representative version of the tasks in the user study, where they were asked to grasp and translate a cube object in the 3D space.

3.5.3 Test

Once participants were comfortable with the interaction space and the overall VR environment, they were presented with the main experimental task. Each participant completed 16 grasps, with a total of 624 grasps recorded during the study (16 grasps × 39 participants). Participants were seated on a chair placed at the centre of the virtual environment to allow them to focus on the hand interaction exclusively. The objects presented in were randomly positioned on the virtual table for each task, but were consistent across participants. Participants were instructed to grasp the virtual object they way they felt most intuitive, notifying the test instructor when they were happy with their grasp.

Hence, the trigger for the interaction was controlled by the test coordinator, thus, allowing the capture of intuitive user grasps that were not constrained by an automatic interaction trigger. The task was completed when all the objects from the virtual table were arranged at their target location and rotation. Hand tracking data was then recorded from the Leap Motion device, while an image was captured of both the VR scenario and of the real hand (from the webcam). The position of the objects on the virtual table was randomised and each participant was presented with the same setting.

4 EXPERIMENT 2 [E2]

Repeat studies are often conducted to confirm a study's findings in conditions similar to the original study [42]. Accordingly, we

(a) Experiment 2 - Condition 1 (E21)   (b) Experiment 2 - Condition 2 (E22)

Figure 6: Experiment 2 conditions: (a) Experiment 2 - Condition 1 [E21] where the target objects position is changed: *Fruits* to the left, *Tools* centered and *Groceries* to the right, with interaction objects rotated at 60°; and (b)Experiment 2 - Condition 2 [E22] where the target object position is changed: *Groceries* to the left, *Fruits* centered and *Tools* to the right, with interaction objects rotated at 90°.

conducted a second experiment consisting of two similar tasks, following the methodology presented in Section 3, to confirm our findings from the first experiment and understand if our approach is repeatable under different conditions:

4.1 Virtual Object Manipulation - Placement:

We change the placement of our target objects based on categories (see Section Target Categories) as follows; In [E21] we position the *Fruits* at the left, *Tools* in the centre and *Groceries* at the right of the virtual scene. [E22] had the *Groceries* at the left, *Fruits* in the centre and *Tools* at the right of the virtual scene.

4.2 Virtual Object Manipulation - Rotation:

We change the rotation of our interaction objects as follows; In [E21] we rotate the virtual objects placed on the virtual table by 60°while in [E22] we rotate them by 90°.

5 METRICS

5.1 Labelling

To label the grasps collected (images from the virtual and real camera) (39 *participants* $\times$ 16 *objects* $\times$ 3 *tasks* = 1872 *grasps*) we followed the methodology of Feix *et al.* [31, 32]. Two academic members of the staff with background in grasping literature were trained to annotate the 1872 grasps collected during the series of experiments. The metrics used for labelling are grasp type (see Section 5.2) and grasp dimension (see Section 5.3) For the first stage, two trained rater academics labelled all grasps individually; following the methodology described in [31, 32]. The raters came from a computer science background and were familiar with human grasping literature. Raters were asked to label images with the real and virtual hand collected during the experiments. Raters were given the option to rate grasps as "cannot classify". The difference in the parameters between raters were analysed by rater 1, who made a final decision about which rater's assignment was correct as in [31, 32].

5.2 Grasp Type

The full set of grasps used for labelling are those by Feix *et al.* in the Human GRASP taxonomy [14]. This taxonomy divides grasp types in three main categories: *Power*, *Intermediate* and *Precision*, with each category then subdivided according to the thumb position into *Thumb Abducted* and *Thumb Adducted*.

Power grasps: *Power* grasps are linked to stability and security. These grasps are distinguished by large areas of contact between the hand and the object [23]. Grasp types under this category are shown in Figure 7.
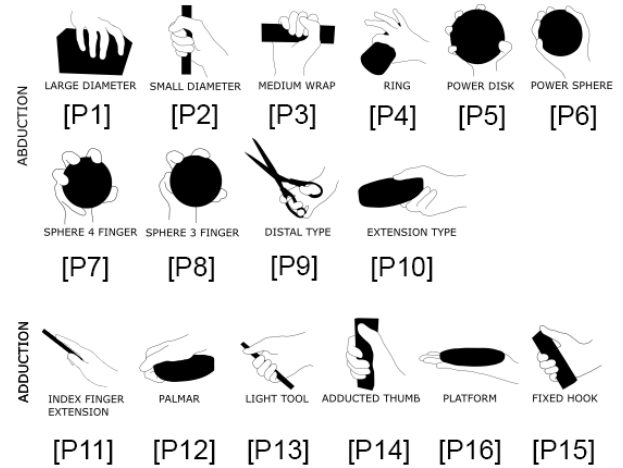

Figure 7: Grasp types from *Power* grasp category, subdivided in *Thumb Adducted* and *Thumb Abducted* as presented in the Human Grasp Taxonomy for real objects [14].

Precision grasps: In these grasps the object is commonly held between the finger tips. While this allows an increased level of manipulation by movement of the fingertips, the object cannot be gripped firmly [49]. The set of *Precision* grasps used for labelling is displayed in Figure 8.

Intermediate grasps: These grasps present elements of *Power* and *Precision* roughly in the same proportion, enabling a finer representation of grasp types [14]. Representative examples of the grasps in this category are displayed in Figure 9.

5.3 Grasp Dimension

Grasp dimension (GDisp) is defined as the part of the object that lies between the fingers when grasped [32]. Three grasp dimensions are defined in the literature: A (longest), B and C(shortest) For each grasp, raters assigned a grasp dimension, based on the object axes.

5.4 Statistical Analysis

The Shapiro-Wilk [66] normality test found the data to be not normally distributed. We test for significance between the conditions and the metrics described using a non parametric Friedman test [35].

6 RESULTS

We collected and labelled a total of 1872 grasps for which we present the results in this section. Figure 10 shows the overall results for each experiment condition [E1], [E21] and [E22] and each object category *Equant*, *Oblate*, *Prolate* and *Bladed*. We present main grasps which show the top three grasps in percentages and number of instances (Grasp types can be found in Figures 7,8 and 9) as well as grasp dimension in percentages and number of instances.

6.1 Most Used Grasp Types

The most used grasps for [E1] and [E2] were: *Large Diameter [P1]* (28.09%, 526 instances), *Small Diameter [P2]* (11.11%, 208 instances), *Medium Wrap [P3]* (10.30%, 193), *Power Sphere [P6]* (9.77%, 183 instances), *Precision Disk [PC10]*, (8.01%, 150 instances) *Thumb 2-Finger [PC4]* (5.34% , 100 instances), *Stick [I5]* (4.54%, 85 instances) and *Thumb 3-Finger [PC5]* (4.27%, 80 instances), with the first four most used grasp types being from the *Power* category, accounting for 59.59% and 1110 instances. This is presented in Figure 12.
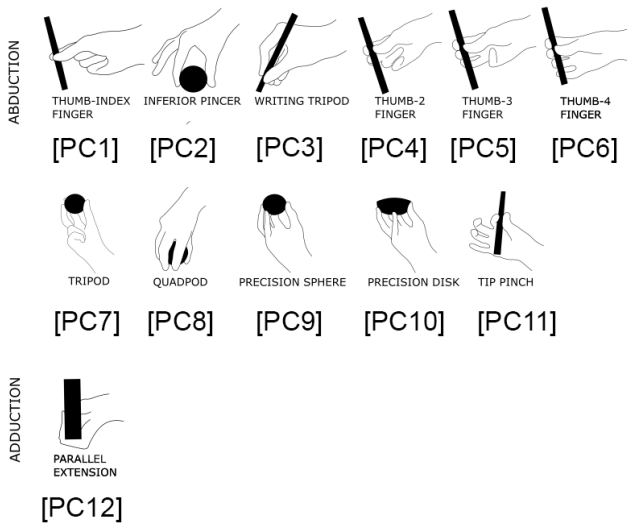
Figure 8: Grasp types from *Precision* grasp category, subdivided in *Thumb Adducted* and *Thumb Abducted* as presented in the Human Grasp Taxonomy for real objects [14]
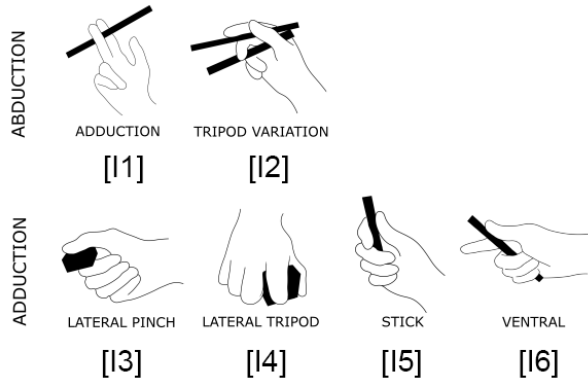

Figure 9: Grasp types from *Intermediate* grasp category, subdivided in *Thumb Adducted* and *Thumb Abducted* as presented in the Human Grasp Taxonomy for real objects [14]

6.2 Grasp Type Categories

Our results show that out of a total of 1872 grasps, 1168 instances (62.39%) were grasped with a *Power* grasp, 614 instances (32.79%) were grasped with *Precision* grasps and 90 instances (4.80%) with *Intermediate* grasps. We statistically analysed the experiment conditions for each object category (Equant [E1], Equant [E21], Equant [E22]) in terms of grasp category and found no significant differences: *Equant* ($\chi\hat{2}$ = 1.41, p-value = 0.49), *Prolate* ($\chi\hat{2}$ = 5.58, p = 0.06) *Oblate* ($\chi\hat{2}$ = 0.03, p = 0.98), *Bladed* ($\chi\hat{2}$ = 0.1827, p = 0.91). Further, we statistically analysed the object categories for each experiment condition (Equant [E1], Prolate [E1], Oblate [E1], Bladed [E1]) in terms of grasp category and found significant differences between object categories for all conditions. [E1] ($\chi\hat{2}$ = 12.31, p 0.01), [E21] ($\chi\hat{2}$ = 30.19, p 0.01), [E22] ($\chi\hat{2}$ = 12.46, p 0.01)

The overall results in percentages and number of instances for each object category are the following: *Equant* category was grasped using *Power* grasps in 71.79% (336 instances), *Precision* grasps in 27.56% (129 instances) and *Intermediate* grasps in 0.64% (3 instances). *Prolate* objects were grasped using *Power* in 58.11% (272

instances), *Precision* in 23.50% (110 instances) and *Intermediate* in 18.37% (86 instances). *Oblate* objects were grasped using *Power* in 64.74% (303 instances), *Precision* in 35.04% (164 instances) and *Intermediate* in 0.21% (1 instance). *Bladed* objects were grasped using *Power* in 54.91% (257 instances), *Precision* in 45.08% (211 instances) and no instances using *Intermediate* grasps. Further, we present overall results for each experiment condition: [E1] showed 59.93% (374 instances) for *Power*, 34.93% (218 instances) for *Precision* and 5.12% (32 instances) for *Intermediate*. [E21] showed 62.01% (387 instances) for *Power*, 31.89% (199 instances) for *Precision* and 6.08% (38 instances) for *Intermediate*. [E22] showed 65.22% (407 instances) for *Power*, 31.57% (197 instances) for *Precision* and 3.20% (20 instances) for *Intermediate*.

6.3 Grasp Dimension

We statistically analysed the experiment conditions for each object category (Equant [E1], Equant [E21], Equant [E22]) in terms of grasp dimension and found no statistical differences between experiment conditions: *Equant* ($\chi\hat{2}$ = 0.73, p = 0.69) *Prolate* ($\chi\hat{2}$= 0.37, p = 0.82), *Oblate* ($\chi\hat{2}$ = 0.87, p = 0.64) *Bladed* ($\chi\hat{2}$ = 5.11, p = 0.07). Further, we statistically analysed the object categories for each experiment condition (Equant [E1], Prolate [E1], Oblate [E1], Bladed [E1]) and found significant differences between object categories in terms of grasp dimensions: [E1] ($\chi\hat{2}$ = 30.10, p ¡ 0.01), [E21] ($\chi\hat{2}$ = 20.59, p ¡ 0.01), [E22] ($\chi\hat{2}$ = 39.67, p ¡ 0.01)

The overall results in terms of grasp dimension for each object characteristics are the following: *Equant* objects were grasped in A (0), B = 16.66% (78 instances) and C = 83.33% (390 instances). *Prolate* objects were grasped in A = 0.21% (1 instance), B = 35.04% (164 instances) and C = 64.74% (303 instances). *Oblate* objects were grasped in A = 1.28% (6 instances), B = 23.93% (112 instances) and C = 74.78% (350 instances).*Bladed* objects were grasped in A = 0.85% (4 instances), B = 53.20% (249 instances) and C = 45.94% (215 instances). Further, we present results for each experiment condition: [E1] showed 68.42% (427 instances) for C dimension, 31.25% (195 instances) for B dimension and 0.32% (2 instances) for A dimension. [E21] showed 70.67% (441 instances) for C dimension, 28.52% (178 instances) for B dimension and 0.80% (5 instances) for A dimension. [E22] showed 61.69% (385 instances) for C dimension, 37.66% (235 instances) for B dimension and 0.64% (4 instances) for A dimension.

7 DISCUSSION

Considerable effort has been made in determining how certain parameters influence human grasping when manipulating real objects [31] and it was found that grasp choice is at least 43% influenced by object properties and 31% influenced by the task properties assigned in the study [31]. Therefore, in this section we discuss how these parameters influence grasp choice when manipulating 3D objects in VR.

7.1 Grasp Types

When analysing grasping real objects, a total of 13 different grasp types accounted for 82.80% of the data [32]. Our results show that in VR only 8 grasps account for 81.34%, suggesting that fewer grasp types are required when manipulating virtual objects, which might be due to the lack of sensory feedback in VR [21] which allows for a less constrained (by the mass and shape of the object) interaction in VR. Additionally, out of the 34 available grasp types, our experiments revealed that only 19 were used. We therefore contribute to the community by showing that more than 80% of the virtual objects could be grasped using only 8 grasp types in VR.

7.2 Grasp Aperture

Grasp aperture, which is the distance between thumb and index finger tips [27], has shown to be influenced by size and shape in real

Figure 10 table:

Ctg.	Obj.	E1 Main Grasps			E1 GDim A	B	C	E2-1 Main Grasps			E2-1 GDim A	B	C	E2-2 Main Grasps			E2-2 GDim A	B	C
Equant	(brick)	[P1] 53.84% [21]	[P6] 10.25% [4]	[PC5] 10.25% [4]	0% [0]	23.07% [9]	76.92% [30]	[P1] 51.28% [20]	[PC5] 12.82% [4]	[P3] 10.25% [4]	0% [0]	30.76% [12]	69.23% [27]	[P1] 61.53% [24]	[P3] 15.38% [6]	[PC4] 7.69% [3]	0% [0]	41.02% [16]	58.97% [23]
	(mug)	[P1] 56.41% [22]	[P2] 25.64% [10]	[P6] 5.12% [2]	0% [0]	20.51% [8]	79.48% [30]	[P1] 41.02% [16]	[P2] 38.46% [15]	[P6] 7.69% [3]	0% [0]	15.38% [6]	84.61% [33]	[P1] 41.02% [16]	[P2] 38.46% [15]	[PC1] 7.69% [3]	0% [0]	23.07% [9]	76.92% [30]
	(orange)	[P6] 76.92% [30]	[PC9] 10.25% [4]	[PC7] 5.12% [2]	0% [0]	5.12% [2]	94.85% [37]	[P6]	[PC9] 17.94% [7]	[PC7] 10.25% [4]	0% [0]	5.12% [2]	94.87% [37]	[P6] 84.61% [33]	[PC9] 10.25% [4]	[P7] 2.56% [1]	0% [0]	7.69% [3]	92.30% [36]
	(cube)	[P6] 28.20% [11]	[PC7] 15.38% [6]	[PC9] 12.82% [5]	0% [0]	5.12% [2]	94.85% [37]	[P6]	[PC9] 17.94% [7]	[PC7] 10.25% [4]	0% [0]	12.82% [5]	87.17% [34]	[P6] 53.84% [21]	[PC8] 12.82% [5]	[PC9] 10.25% [4]	0% [0]	10.25% [4]	89.74% [35]
Prolate	(banana)	[P3] 74.35% [29]	[PC4] 7.69% [3]	[PC5] 5.12% [2]	0% [0]	7.69% [3]	92.30% [36]	[P3]	[P2] 17.94% [7]	[PC4] 10.25% [4]	0% [0]	10.25% [4]	89.74% [35]	[P3] 56.41% [22]	[P2] 25.64% [10]	[PC4] 10.25% [4]	0% [0]	7.69% [3]	92.30% [36]
	(rolling pin)	[P3] 46.15% [18]	[P2] 33.33% [13]	[PC4] 7.69% [3]	0% [0]	87.17% [34]	12.82% [5]	[P3]	[P2] 25.64% [10]	[PC4] 10.25% [4]	0% [0]	69.23% [27]	30.76% [12]	[P3] 46.15% [18]	[P2] 23.07% [9]	[PC4] 10.25% [4]	2.56% [1]	69.23% [27]	28.20% [11]
	(flashlight)	[I5] 46.15% [18]	[P2] 17.94% [7]	[P13] 10.25% [4]	0% [0]	28.20% [11]	71.79% [28]	[I5]	[P2] 25.64% [10]	[PC4] 10.25% [4]	0% [0]	33.33% [13]	66.66% [26]	[I5] 35.89% [14]	[P2] 33.33% [13]	[PC4] 7.69% [3]	0% [0]	20.51% [8]	79.48% [31]
	(brush)	[P2] 25.64% [10]	[I5] 25.64% [10]	[PC6] 20.51% [8]	0% [0]	20.51% [8]	79.48% [31]	[I5]	[P2] 20.51% [8]	[P13] 17.94% [7]	0% [0]	17.94% [7]	82.05% [32]	[P2] 56.41% [22]	[PC4] 15.38% [6]	[I5] 10.25% [4]	0% [0]	48.71% [19]	51.28% [20]
Oblate	(cracker box)	[P1] 69.23% [27]	[PC4] 7.69% [3]	[PC5] 7.69% [3]	0% [0]	0% [0]	100% [39]	[P1]	[P2] 12.82% [5]	[PC5] 7.69% [3]	0% [0]	0% [0]	100% [39]	[P1] 71.79% [28]	[PC5] 10.25% [4]	[P3] 7.69% [3]	0% [0]	15.38% [6]	84.61% [33]
	(glove)	[PC7] 20.51% [8]	[PC10] 17.94% [7]	[P5] 15.38% [6]	0% [0]	87.27% [34]	12.82% [5]	[P5]	[PC10] 15.38% [6]	[P6] 12.82% [5]	2.56% [1]	76.92% [30]	20.51% [8]	[PC10] 20.51% [8]	[P5] 17.94% [7]	[PC1] 12.82% [5]	0% [0]	87.17% [34]	12.82% [5]
	(jello)	[P1] 20.51% [8]	[P3] 25.64% [10]	[PC4] 17.94% [7]	5.12% [2]	5.12% [2]	89.74% [35]	[P1] 38.46% [15]	[P3] 28.20% [11]	[PC4] 15.38% [6]	2.56% [1]	2.56% [1]	94.87% [37]	[P1] 43.58% [17]	[P3] 20.51% [8]	[PC4] 12.82% [5]	2.56% [1]	2.56% [1]	94.87% [37]
	(spam can)	[P1] 84.61% [33]	[PC4] 5.12% [2]	[P3] 5.12% [2]	0% [0]	2.56% [1]	97.43% [38]	[P1] 76.92% [30]	[PC4] 7.69% [3]	[PC5] 5.12% [2]	2.56% [1]	0% [0]	97.43% [38]	[P1] 69.23% [27]	[PC5] 10.25% [4]	[P6] 5.12% [2]	0% [0]	20.51% [8]	79.48% [31]
Bladed	(mustard)	[P1] 82.05% [32]	[PC4] 5.12% [2]	[P3] 5.12% [2]	0% [0]	12.82% [5]	87.17% [34]	[P1] 84.61% [33]	[PC5] 5.12% [2]	[P3] 5.12% [2]	2.56% [1]	5.12% [2]	92.30% [36]	[P1] 79.48% [31]	[PC4] 7.69% [3]	[PC5] 7.69% [3]	0% [0]	7.69% [3]	92.30% [36]
	(knife)	[PC10] 46.15% [18]	[P2] 10.25% [4]	[PC4] 10.25% [4]	0% [0]	82.05% [32]	17.94% [7]	[PC10] 46.15% [18]	[PC4] 12.82% [5]	[PC7] 10.25% [4]	0% [0]	97.43% [38]	2.56% [1]	[PC10] 38.46% [15]	[PC4] 15.38% [6]	[P5] 12.82% [5]	5.12% [2]	87.17% [34]	7.69% [3]
	(spray bottle)	[P1] 84.61% [33]	[PC4] 7.69% [3]	[P3] 5.12% [2]	0% [0]	33.33% [13]	66.66% [26]	[P1] 82.05% [32]	[PC5] 5.12% [2]	[P3] 5.12% [2]	0% [0]	0% [0]	100% [39]	[P1] 76.92% [30]	[PC4] 7.69% [3]	[P6] 7.69% [3]	0% [0]	58.97% [23]	41.02% [16]
	(tray)	[PC10] 64.10% [25]	[P2] 7.69% [3]	[PC4] 7.69% [3]	0% [0]	79.48% [31]	20.51% [8]	[PC10] 48.71% [19]	[P5] 17.94% [7]	[P2] 12.82% [5]	2.56% [1]	79.48% [31]	17.94% [7]	[PC10] 66.66% [26]	[P5] 10.25% [4]	[P2] 7.69% [3]	0% [0]	94.87% [37]	5.12% [2]

Legend: ■ Power Grasps ▬ Precision Grasps ▬ Intermediate Grasps

Figure 10: Results showing top three grasps used (Main Grasps), grasp dimension (GDim) for each object category (Ctg), each object type (Obj) and each task from [E1], [E21] and [E22]. For each grasp type we show the grasp code (please see Figures presented in Section 5.2), the percentage and the number of instances. Grasp types are colour coded based on their category: *Power*, *Precision* and *Intermediate*. Grasp dimension is shown as percentages and number of instances for each grasping dimension available (A, B or C)

object grasping [31].

Studies looking at grasping patterns in real environments showed that the *Medium Wrap [P3]* grasp type from the *Power* grasp category is the most common grasp used when manipulating real objects [31]. However, our results show that when grasping virtual objects for docking tasks the most used grasp type is *Large Diameter [P1]* from the *Power* grasp category. Comparing the two grasp types, *Large Diameter [P1]* is characterised by a larger grasp aperture than

Medium Wrap [P3], therefore suggesting that users grasp virtual objects with a larger grasp than they grasping real objects. This finding correlates to previous work that showed that hand aperture is wider when grasping objects in a virtual environment compared to a physical environment [51]. Moreover, this could also imply that as opposed to grasping real objects when the shape of the hand evolves gradually to conform the contours of the object [62], when grasping virtual objects users focused less on the contours of the
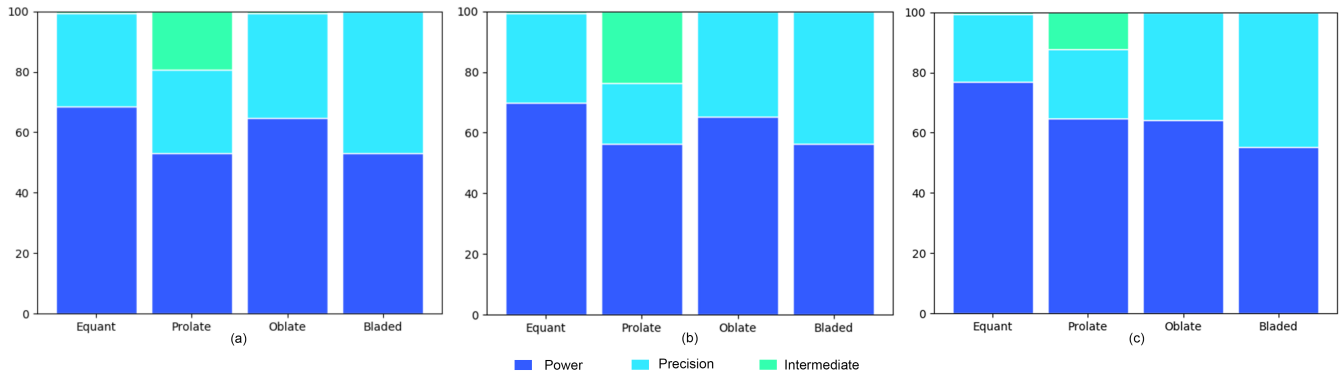
Figure 11: Grasp categories *(Power, Precision, Intermediate)* percentages based on object categories *(Equant, Prolate, Oblate, Bladed)* for (a) Experiment 1 [E1], (b) Experiment 2 - Condition 1 [E21] and (c) Experiment 2 - Condition 2 [E22].

objects which might be due to the lack of haptic feedback.

In real object interaction, the grasping of an object depends on different conditions, such as: the behaviour of the object under the influence of gravity and the surface material of the object and the geometrical conditions at the point of contact between the object and the grasping hand [70].

7.3 Object Categories

While this methodology has not been applied before for virtual objects, we applied Zingg's [76] methodology to categorise objects based on their virtual shape. We then compared these categories across grasping metrics for each experiment condition and show that there are significant differences between object categories for grasp metrics, showing that virtual object shape influences grasping approach in VR. A similar result was observed in Mixed Reality, namely that object shape and size influenced the grasping approach [2]. Further, we compared experiment conditions for each object category and found that there were no differences between conditions for grasp metrics. This shows that grasp patterns for each object category was consistent for different positions and orientations. Our results along with this method for categorising virtual objects can then be taken forward to assess other virtual objects and parameterize our findings for achieving a natural and intuitive grasp model for interactions in VR.

7.4 Grasp Dimension

We show that users predominantly grasped the virtual objects along the C dimension which is the shortest, followed by the B dimension, with few or no instances presented for A dimension. When comparing object categories for grasp dimension, significant differences were observed for all experimental conditions, showing that object shape had an influence on grasp dimension.This result

Figure 12: Overall most used used grasp types for Experiment 1 [E1], Experiment 2 - Condition 1 [E21] and Experiment 2 - Condition 2 [E22] shown in percentages.

is consistent with real object grasping research [32], showing that there are takeaways that users take when grasping virtual objects from real life experience of grasping objects. However, different orientation or placement of the object did not show any difference in terms of grasping dimension, which confirms findings from real object grasping literature, showing that the grasped dimension is defined in conjunction with the object dimensions, inherently linked to object properties [22]. While *Equant*, *Prolate* and *Oblate* presented a higher number of instances for C dimension, *Bladed* objects showed a more balanced approach, with both B and C dimensions being preferred. This might be due to *Bladed* objects presenting a long and thin shape, with both B and C being easily grasped by the human hand. This is consistent with real object grasping literature, where long objects are grasped along B/C dimension [32].

7.5 Docking Tasks

Our results show that grasp metrics did not change for different orientations and positions of the virtual objects. This is consistent with real object grasping research showing that grasping is highly influenced by the task to be performed [31].

7.6 Recommendations

Current limitations in equipment lead to physical forces missing in most immersive scenarios [38], which has shown to induce a shift from natural towards pantomimed grasps [26]. This has shown to affect the naturalness of interaction and forcing users to change their approach when grasping objects in VR [12]. Our experiments provide an insight into how users intuitively grasp virtual objects, presenting grasp patterns and how they are influenced by object characteristics and task properties. To understand grasping patterns in VR we categorise virtual objects using Zingg's [76] methodology for categorising real objects, which has not been previously done in VR. This methodology could be further replicated for other virtual objects to further investigate interaction patterns and how they are influenced by object characteristics.

Further, we present grasping patterns based on object category and task properties, showing that grasp type metric shows that users grasp virtual objects differently than they grasp real objects, while grasp dimension shows that users interacted with the virtual objects using the knowledge from real object manipulation. Therefore, a natural grasp model based on our findings would parameterize the grasp types used for interacting, however the grasp dimension will remain the same as for real object grasping. Our grasping patterns are useful for designers creating VR environments that require direct interaction with objects, specially for training (i.e. VR training environments in construction and manufacturing [5], surgical training techniques [54]) where the user experience could
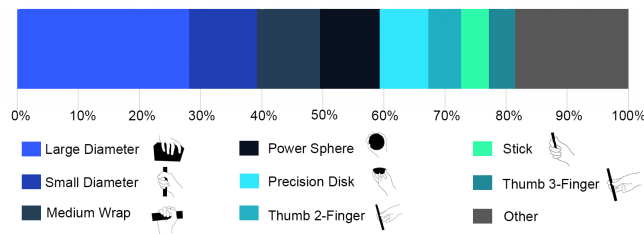
be vulnerable [17]. In these scenarios, the shape of the objects needs to be considered, with objects being designed to specifically encourage this user selection of precision grasps and thus facilitating a more precise and finer manipulation. Therefore, categorising the object and designing the interactions accordingly would encourage the design of more natural interactions. Additionally, our results can be further used to parameterize grasp types and develop natural and intuitive grasp models, where each of the grasps here can be defined by core parameters relating to hand tracking sensor data. This alongside Zingg's methods for categorizing object type, will lead toward a more natural and reliable interaction for freehand grasping.

8 CONCLUSION

We conducted two elicitation experiments to understand how task properties and object characteristics influence grasp patterns in VR. For this, we asked users to grasp 16 virtual objects categorised by shape using Zingg's [76] methodology, in a mixed docking task exploring rotation, placement and target location. The virtual objects employed are representations of real objects and thus used as referents of intuitive grasping from real life. We followed a *Wizard of Oz* methodology for freehand grasping extracting grasp types, grasp category and grasp dimension. We show grasping patterns for each experiment condition and object category and discuss their implications for 3D interaction design. We found that only 8 grasps account for more than 80% of the data when grasping in VR, which shows users' choice for grasp types vary less in VR compared to real object manipulation. Moreover, we show that users grasp virtual objects with a larger grasp aperture than real objects. However, while grasp type metric showed a difference in grasping approach in VR, grasp dimension showed that there are takeaways from real life experience that users take when grasping virtual objects, namely grasping objects on the shortest dimension. We then provided recommendations VR interaction design, showing that our results can be furthered to apply this categorisation methodology for other virtual objects when assessing interaction paradigms, and parameterize the identified grasp types for development of a natural and intuitive grasp models.

REFERENCES

[1] V. K. Adhikarla, P. Woźniak, A. Barsi, D. Singhal, P. T. Kovács, and T. Balogh. Freehand interaction with large-scale 3d map data. In *2014 3DTV-Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON)*, pp. 1–4, July 2014. doi: 10.1109/3DTV.2014.6874711

[2] M. Al-Kalbani, I. Williams, and M. Frutos-Pascual. Analysis of medium wrap freehand virtual object grasping in exocentric mixed reality. In *Mixed and Augmented Reality (ISMAR), 2016 IEEE International Symposium on*, pp. 84–93. IEEE, 2016.

[3] H. Bai, L. Gao, J. El-Sana, and M. Billinghurst. Markerless 3D gesture-based interaction for handheld Augmented Reality interfaces. *2013 IEEE International Symposium on Mixed and Augmented Reality, ISMAR 2013*, (October 2013):0–5, 2013.

[4] H. Bai, A. Nassani, B. Ens, and M. Billinghurst. Asymmetric bimanual interaction for mobile virtual reality. In *Proceedings of the 27th International Conference on Artificial Reality and Telexistence and 22nd Eurographics Symposium on Virtual Environments*, ICAT-EGVE '17, p. 83–86. Eurographics Association, Goslar, DEU, 2017.

[5] R. Barkokebas, C. Ritter, V. Sirbu, X. Li, and M. Al-Hussein. Application of virtual reality in task training in the construction manufacturing industry. In *ISARC. Proceedings of the International Symposium on Automation and Robotics in Construction*, vol. 36, pp. 796–803. IAARC Publications, 2019.

[6] F. Biocca. Communication applications of virtual reality. 06 1995.

[7] A. D. Blaga, M. Frutos-Pascual, M. Al-Kalbani, and I. Williams. [poster] usability analysis of an off-the-shelf hand posture estimation sensor for freehand physical interaction in egocentric mixed reality. In *2017 IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct)*, pp. 31–34, Oct 2017.

[8] A. D. Blaga, M. Frutos-Pascual, C. Creed, and I. Williams. Too hot to handle: An evaluation of the effect of thermal visual representation on user grasping interaction in virtual reality. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20, p. 1–16. Association for Computing Machinery, New York, NY, USA, 2020. doi: 10.1145/3313831.3376554

[9] J. Boritz and K. Booth. A study of interactive 6 dof docking in a computerised virtual environment. *Proceedings. IEEE 1998 Virtual Reality Annual International Symposium (Cat. No.98CB36180)*, pp. 139–146, 1998.

[10] C. W. Borst and A. P. Indugula. Realistic virtual grasping. In *IEEE Proceedings. VR 2005. Virtual Reality, 2005.*, pp. 91–98, 2005.

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

[12] C. W. Borst and M. Prachyabrued. Nonuniform and adaptive coupling stiffness for virtual grasping. In *Virtual Reality (VR), 2013 IEEE*, pp. 35–38. IEEE, 2013.

[13] I. M. Bullock, T. Feix, and A. M. Dollar. Finding small, versatile sets of human grasps to span common objects. In *2013 IEEE International Conference on Robotics and Automation*, pp. 1068–1075, May 2013. doi: 10.1109/ICRA.2013.6630705

[14] I. M. Bullock, J. Z. Zheng, S. D. L. Rosa, C. Guertler, and A. M. Dollar. Grasp frequency and usage in daily household and machine shop tasks. *IEEE Transactions on Haptics*, 6(3):296–308, July 2013. doi: 10.1109/TOH.2013.6

[15] B. Calli, A. Singh, A. Walsman, S. Srinivasa, P. Abbeel, and A. M. Dollar. The ycb object and model set: Towards common benchmarks for manipulation research. In *2015 International Conference on Advanced Robotics (ICAR)*, pp. 510–517, 2015. doi: 10.1109/ICAR.2015.7251504

[16] B. Calli, A. Walsman, A. Singh, S. Srinivasa, P. Abbeel, and A. M. Dollar. Benchmarking in manipulation research: Using the yale-cmu-berkeley object and model set. *IEEE Robotics Automation Magazine*, 22(3):36–52, Sep. 2015. doi: 10.1109/MRA.2015.2448951

[17] T. Chen, Y.-S. Wu, and K. Zhu. Investigating different modalities of directional cues for multi-task visual-searching scenario in virtual reality. In *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology*, VRST '18. Association for Computing Machinery, New York, NY, USA, 2018. doi: 10.1145/3281505.3281516

[18] E. Choi, S. Kwon, D. Lee, H. Lee, and M. K. Chung. Can user-derived gesture be considered as the best gesture for a command?: Focusing on the commands for smart home system. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 56(1):1253–1257, 2012. doi: 10.1177/1071181312561222

[19] I. Choi, E. W. Hawkes, D. L. Christensen, C. J. Ploch, and S. Follmer. Wolverine: A wearable haptic interface for grasping in virtual reality. In *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 986–993. IEEE, 2016.

[20] P. Cohen. The pragmatics of referring and the modality of communication. *Computational Linguistics*, 10:97–146, 01 1984.

[21] N. Cooper, F. Milella, C. Pinto, I. Cant, M. White, and G. Meyer. The effects of substitute multisensory feedback on task performance and the sense of presence in a virtual reality environment. *PloS one*, 13(2):e0191846, 2018.

[22] R. Cuijpers, J. Smeets, and E. Brenner. On the relation between object shape and grasping kinematics. *Journal of neurophysiology*, 91:2598–606, 07 2004. doi: 10.1152/jn.00644.2003

[23] M. R. Cutkosky. On grasp choice, grasp models, and the design of hands for manufacturing tasks. *IEEE Transactions on Robotics and Automation*, 5(3):269–279, Jun 1989. doi: 10.1109/70.34763

[24] N. Dahlbäck, A. Jönsson, and L. Ahrenberg. Wizard of oz studies: Why and how. In *Proceedings of the 1st International Conference on Intelligent User Interfaces*, IUI '93, p. 193–200. Association for Computing Machinery, New York, NY, USA, 1993. doi: 10.1145/169891.169968

[25] D. Datcu and S. Lukosch. Free-hands interaction in augmented reality. In *Proceedings of the 1st Symposium on Spatial User Interaction*, SUI

'13, pp. 33–40. ACM, 2013. doi: 10.1145/2491367.2491370

[26] A. David Milner, T. Ganel, and M. Goodale. Does grasping in patient d.f. depend on vision? *Trends in cognitive sciences*, 16:256–7; discussion 258, 03 2012. doi: 10.1016/j.tics.2012.03.004

[27] A. Edsinger and C. C. Kemp. Human-robot interaction for cooperative manipulation: Handing objects to one another. *RO-MAN 2007 - The 16th IEEE International Symposium on Robot and Human Interactive Communication*, pp. 1167–1172, 2007.

[28] J. Eisenstein and R. Davis. Visual and linguistic information in gesture classification. *SIGGRAPH 2006 - ACM SIGGRAPH 2006 Courses*, 2006. doi: 10.1145/1185657.1185791

[29] D. El-Shimy, G. Marentakis, and J. Cooperstock. Tech-note: Multimodal feedback in 3d target acquisition. pp. 95–98, 01 2009. doi: 10.1109/3DUI.2009.4811212

[30] D. Englmeier, J. Dörner, A. Butz, and T. Höllerer. A tangible spherical proxy for object manipulation in augmented reality. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 221–229, 2020.

[31] T. Feix, I. M. Bullock, and A. M. Dollar. Analysis of Human Grasping Behavior: Correlating Tasks, Objects and Grasps. *IEEE Transactions on Haptics (in press)*, 7(3):311–323, 2014.

[32] T. Feix, I. M. Bullock, and A. M. Dollar. Analysis of human grasping behavior: Object characteristics and grasp type. *IEEE Transactions on Haptics*, 7(3):311–323, July 2014. doi: 10.1109/TOH.2014.2326871

[33] T. Feix, R. Pawlik, H. B. Schmiedmayer, J. Romero, and D. Kragi. A comprehensive grasp taxonomy. In *Robotics Sci-ence and Systems Conference: Workshop on Understanding the Human Hand for Advancing Robotic Manipulation*, pp. 2–3, 2009.

[34] T. Feix, J. Romero, H.-B. Schmiedmayer, A. M. Dollar, and D. Kragiç. The grasp taxonomy of human grasp types. *IEEE Transactions on Human-Machine Systems*, 46(1):66–77, Feb 2016. doi: 10.1109/THMS.2015.2470657

[35] M. Friedman. A comparison of alternative tests of significance for the problem of m rankings. *The Annals of Mathematical Statistics*, 11(1):86–92, 1940.

[36] M. Frutos-Pascual, J. M. Harrison, C. Creed, and I. Williams. Evaluation of ultrasound haptics as a supplementary feedback cue for grasping in virtual environments. In *2019 International Conference on Multimodal Interaction*, pp. 310–318, 2019.

[37] K. Hinckley, R. Pausch, J. C. Goble, and N. F. Kassell. Passive real-world interface props for neurosurgical visualization. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '94, p. 452–458. Association for Computing Machinery, New York, NY, USA, 1994. doi: 10.1145/191666.191821

[38] D. Holz, S. Ullrich, M. Wolter, and T. Kuhlen. Multi-contact grasp interaction for virtual environments. *JVRB - Journal of Virtual Reality and Broadcasting*, 5(2008)(7), 2008.

[39] S. Hudson, S. Matson-Barkat, N. Pallamin, and G. Jegou. With or without you? interaction and immersion in a virtual reality experience. *Journal of Business Research*, 2018. doi: 10.1016/j.jbusres.2018.10.062

[40] Y. Jia and A. G. Campbell. Immersive virtual reality training tool for iot device placement. In *Proceedings of the 10th EAI International Conference on Simulation Tools and Techniques*, SIMUTOOLS '17, pp. 81–86, 2017. doi: 10.1145/3173519.3173521

[41] X. Jiang, Z. Xiao, and C. Menon. Virtual grasps recognition using fusion of leap motion and force myography. *Virtual Reality*, 03 2018. doi: 10.1007/s10055-018-0339-2

[42] K. Jones, P. Derby, and L. Schmidlin. An investigation of the prevalence of replication research in human factors. *Human factors*, 52:586–95, 10 2010. doi: 10.1177/0018720810384394

[43] K. Kilteni, R. Groten, and M. Slater. The sense of embodiment in virtual reality. *Presence Teleoperators amp Virtual Environments*, 21, 11 2012. doi: 10.1162/PRES_a_00124

[44] F. Kistler and E. André. User-defined body gestures for an interactive storytelling scenario. In *INTERACT*, 2013.

[45] C. Kray, D. Nesbitt, J. Dawson, and M. Rohs. User-defined gestures for connecting mobile phones, public displays, and tabletops. In *Proceedings of the 12th International Conference on Human Computer Interaction with Mobile Devices and Services*, MobileHCI '10, p. 239–248. Association for Computing Machinery, New York, NY, USA, 2010. doi: 10.1145/1851600.1851640

[46] P. Krompiec and K. Park. Enhanced player interaction using motion controllers for first-person shooting games in virtual reality. *IEEE Access*, 7:124548–124557, 2019. doi: 10.1109/ACCESS.2019.2937937

[47] G. Lawson, D. Salanitri, and B. Waterfield. Future directions for the development of virtual reality within an automotive manufacturer. *Applied Ergonomics*, 53:323 – 330, 2016. Transport in the 21st Century: The Application of Human Factors to Future User Needs.

[48] L. Lin and S. Jörg. Need a hand?: How appearance affects the virtual hand illusion. In *Proceedings of the ACM Symposium on Applied Perception*, SAP '16, pp. 69–76, 2016. doi: 10.1145/2931002.2931006

[49] D. Lyons. A simple set of grasps for a dextrous hand. In *Robotics and Automation. Proceedings. 1985 IEEE International Conference*, vol. 2, pp. 588 – 593, 1985.

[50] S. Ma, M. Varley, L. Shark, and J. Richards. Emg biofeedback based vr system for hand rotation and grasping rehabilitation. In *2010 14th International Conference Information Visualisation*, pp. 479–484, 2010. doi: 10.1109/IV.2010.73

[51] E. Magdalon, S. Michaelsen, A. Quevedo, and M. Levin. Comparison of grasping movements made by healthy subjects in a 3-dimensional immersive virtual versus physical environment. *Acta psychologica*, 138:126–34, 06 2011. doi: 10.1016/j.actpsy.2011.05.015

[52] A. Martinet, G. Casiez, and L. Grisoni. The effect of dof separation in 3d manipulation tasks with multi-touch displays. *Proceedings of VRST'10, the 17th ACM Symposium on Virtual Reality Software and Technology*, 01 2010. doi: 10.1145/1889863.1889888

[53] H. Mousavi Hondori, M. Khademi, L. Dodakian, S. Cramer, and C. Lopes. A spatial augmented reality rehab system for post-stroke hand rehabilitation. *Studies in health technology and informatics*, 184:279–85, 02 2013. doi: 10.3233/978-1-61499-209-7-279

[54] Z. H. Nayer, B. Murdock, I. P. Dharia, and D. A. Belyea. Predictive and construct validity of virtual reality cataract surgery simulators. *Journal of Cataract & Refractive Surgery*, 46(6):907–912, 2020.

[55] W. Park. A multi-touch gesture vocabulary design methodolgy for mobile devices. *Division of Mechanical and In-dustrial Engineering POSTECH*, 2012.

[56] A. Peer and K. Ponto. Evaluating perceived distance measures in roomscale spaces using consumer-grade head mounted displays. In *2017 IEEE Symposium on 3D User Interfaces (3DUI)*, pp. 83–86. IEEE, 2017.

[57] C. Peng, J. T. Hansberger, L. Cao, and V. A. Shanthakumar. Hand gesture controls for image categorization in immersive virtual environments. In *2017 IEEE Virtual Reality (VR)*, pp. 331–332, March 2017.

[58] T. Pfeiffer. Using virtual reality technology in linguistic research. In *2012 IEEE Virtual Reality Workshops (VRW)*, pp. 83–84, March 2012. doi: 10.1109/VR.2012.6180893

[59] T. Piumsomboon, A. Clark, M. Billinghurst, and A. Cockburn. User-defined gestures for augmented reality. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems*, pp. 955–960. ACM, 2013.

[60] M. Prachyabrued and C. W. Borst. Design and evaluation of visual interpenetration cues in virtual grasping. *IEEE Transactions on Visualization and Computer Graphics*, 22(6):1718–1731, 2016.

[61] B. Redmond, R. Aina, T. Gorti, and B. Hannaford. Haptic characteristics of some activities of daily living. In *Proceedings of the 2010 IEEE Haptics Symposium*, HAPTIC '10, pp. 71–76. IEEE Computer Society, 2010. doi: 10.1109/HAPTIC.2010.5444674

[62] M. Santello, M. Flanders, and J. F. Soechting. Postural hand synergies for tool use. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 18:10105–15, 12 1998. doi: 10.1523/JNEUROSCI.18-23-10105.1998

[63] B. S. Santos, J. Cardoso, B. Q. Ferreira, C. Ferreira, and P. Dias. Developing 3d freehand gesture-based interaction methods for virtual walkthroughs: Using an iterative approach. In *Handbook of Research on Human-Computer Interfaces, Developments, and Applications*, pp. 52–72. IGI Global, 2016.

[64] J. Schild, J. J. Laviola, and M. Masuch. Altering gameplay behavior using stereoscopic 3d vision-based video game design, 2014.

[65] V. Schwind, P. Knierim, C. Tasci, P. Franczak, N. Haas, and N. Henze.

"these are not my hands!": Effect of gender on the perception of avatar hands in virtual reality. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, pp. 1577–1582. ACM, New York, NY, USA, 2017. doi: 10.1145/3025453.3025602

[66] S. S. Shapiro and M. B. Wilk. An analysis of variance test for normality (complete samples). *Biometrika*, 52(3/4):591–611, 1965.

[67] X. SUN, H. LIU, Y. TIAN, G. WU, and Y. GAO. Team effectiveness evaluation and virtual reality scenario mapping model for helicopter emergency rescue. *Chinese Journal of Aeronautics*, 2020. doi: 10.1016/j.cja.2020.06.003

[68] C. Swindells, K. Inkpen, J. Dill, and M. Tory. That one there! pointing to establish device identity. 09 2002. doi: 10.1145/571985.572007

[69] H. Tian, C. Wang, D. Manocha, and X. Zhang. Realtime hand-object interaction using learned grasp space for virtual environments. *IEEE Transactions on Visualization and Computer Graphics*, pp. 1–1, 2018.

[70] T. Ullmann and J. Sauer. Intuitive virtual grasping for non haptic environments. In *Proceedings the Eighth Pacific Conference on Computer Graphics and Applications*, pp. 373–457. IEEE, 2000.

[71] V. Vuibert, W. Stuerzlinger, and J. R. Cooperstock. Evaluation of docking task performance using mid-air interaction techniques. In *Proceedings of the 3rd ACM Symposium on Spatial User Interaction*, SUI '15, p. 44–52. Association for Computing Machinery, New York, NY, USA, 2015. doi: 10.1145/2788940.2788950

[72] C. Ware and R. Arsenault. Frames of reference in virtual object rotation. In *Proceedings of the 1st Symposium on Applied Perception in Graphics and Visualization*, APGV '04, p. 135–141. Association for Computing Machinery, New York, NY, USA, 2004. doi: 10.1145/1012551.1012576

[73] J. O. Wobbrock, H. H. Aung, B. Rothrock, and B. A. Myers. Maximizing the guessability of symbolic input. In *CHI'05 extended abstracts on Human Factors in Computing Systems*, pp. 1869–1872. ACM, 2005.

[74] J. O. Wobbrock, M. R. Morris, and A. D. Wilson. User-defined gestures for surface computing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 1083–1092. ACM, 2009.

[75] J. O. Wobbrock, M. R. Morris, and A. D. Wilson. User-defined gestures for surface computing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '09, p. 1083–1092. Association for Computing Machinery, New York, NY, USA, 2009. doi: 10.1145/1518701.1518866

[76] T. Zingg. *Beitrag zur schotteranalyse*. PhD thesis, ETH Zürich,Zürich, Switzerland, 1935.