Bimanual, mid-air multi-selection techniques in virtual reality

Pantea Habibi[*] Debaleena Chattopadhyay[†]

Department of Computer Science, University of Illinois Chicago

Abstract

Data exploration in virtual reality (VR) often requires multi-selection of objects, which can vary in number, geometry, and density. Unlike device-based multi-selection, barehand techniques remain little explored. In this paper, we present an empirical comparison of four mid-air multi-selection techniques spanning dominant and non-dominant hand actions (n = 20). The symmetric-synchronous and high-tension asymmetric-asynchronous bimanual techniques were not significantly different in terms of efficiency, accuracy, and throughput for within-reach targets. Furthermore, there were no significant differences between low and high-tension asymmetric-asynchronous bimanual techniques. The symmetric-synchronous bimanual technique was most preferred.

Index Terms: Bimanual interaction, mid-air interaction, virtual reality, multi-selection.

1 Introduction

Data exploration in virtual reality (VR) leverages immersive environments to enhance comprehension and manipulation of complex datasets, providing unique opportunities for intuitive interactions [9]. While multiselection has become increasingly relevant in VR data exploration tasks, such as exploring subgraphs, the potential of mid-air interactions—particularly bimanual techniques that facilitate more natural user input for multiselection—are less explored.

Bimanual interaction can be categorized based on whether the roles of the hands are symmetric or asymmetric. Symmetric interactions occur when both hands contribute equally to a task, with interchangeable roles, whereas asymmetric interactions involve different sub-tasks for the dominant and non-dominant hands [4]. These interactions can be synchronous, where both hands move simultaneously, or asynchronous, where they do not [13]. Bimanual interaction has been adopted for various input modalities such as keyboards [7], touch devices [14], and augmented reality [3].

Another aspect of bimanual interaction is bimanual phrasing, where interactions of the right hand are combined with those of the left hand. In gestural interaction, making kinesthetic tension is a way to create a phrase. These kinesthetic tensions tie together a series of inputs into a cohesive gestural phrase that is comprehensible to the user. One method utilizes kinesthetic tension and lateral preference to phrase bimanual interactions consistent with Guiard's principles [4]. This approach involves asynchronous and asymmetric interactions where the right hand's actions depend on the left hand, and the left hand initiates the phrase by creating tension.

Buxton first introduced the concept of applying phrasing for graphical manipulation on computer interfaces, emphasizing muscular tension and kinesthetic continuity to aggregate tasks into single gestures[2]. This concept has been explored in various HCI interactions, such as Leganchuk et al.'s study on bimanual interaction, which highlighted the benefits of two-handed input[5]. Their results showed significantly faster performance with bimanual techniques and suggested that a user interface design should support an appropriate level of chunking to phrases. Additionally, phrasing has been applied to the use of different input devices in combination, such as touch and pen[14].

In touchless interaction, phrasing has been used for bare-hand mode switching [12]. However, there is still a gap in understanding the effect of phrasing on different tasks, such as multi-selection. Multiple object selection (MOS) in VR is understudied, often using serialized single-object selection techniques. Stenholt [10] adapted brush, lasso, and magic wand techniques from 2D to 3D, highlighting performance differences based on object configuration. Montano-Murillo [8] used slicing volumes to select and refine objects in a 3D environment, showing improved accuracy with physical tablet feedback. Stuerzlinger and Smith [11] explored automatic grouping in 3D scenes, though not suitable for all object patterns. All these works are primarily designed for distant objects.

We introduce three bimanual mid-air interaction techniques for multi-selection tasks. These techniques have varied levels of synchronization, symmetry and tension between the two hands. This helps us understand how lateral asymmetry and different configurations influence user performance and preferences in bimanual mid-air interaction techniques for multiselection, which we investigate in a controlled user study. The interactions we designed include Multi-single-selection (Baseline), where both right (R) and left (L) hands can select an individual object, Sphere-SS (symmetric synchronous), which involves performing a gesture with both hands to create a sphere with the distance between pinches as the diameter and the center as the midpoint of the two pinch gestures, and Sphere-AA (asymmetric asynchronous), where the left hand sets the center of the sphere and the right hand adjusts the radius. For Sphere-AA, we included low tension, allowing movement of the selection center until the right hand starts, and high tension, maintaining the tension until the right hand finishes.

We tested the following hypotheses:

H1. Users will perform better multi-selection with bimanual sphere-based methods than with single-select (baseline).

H2. Users will perform better multi-selection with synchronous methods than with asynchronous methods because synchronization allows both hands to perform simultaneously without waiting for the other to complete a step.

H3. Users will perform more accurate multi-selection with bimanual methods that allow less flexibility in muscle tension (high) than when sub-interactions are more strictly controlled (low) tension.

The results of this study will inform future design and research in mixed reality interaction. By understanding these dynamics, we can further optimize VR environments for data exploration, enhancing both the efficiency and intuitiveness of interactions.

2 Method

Participants A total of 20 right-handed participants (11 males, 7 females, 1 non-binary, Mage = 28.6, SD = 4.22, SE = 0.94) were recruited. Approximately 55% of participants had prior experience with mid-air inputs, primarily for gaming, with controllers being the most common device, used by 7 participants.

Apparatus An Oculus Quest 2 VR headset was used. A Windows 10 machine (3.6GHz Intel i7 CPU, GeForce GTX 1080

[*]e-mail: phabib4@uic.edu

[†]e-mail:debchatt@uic.edu

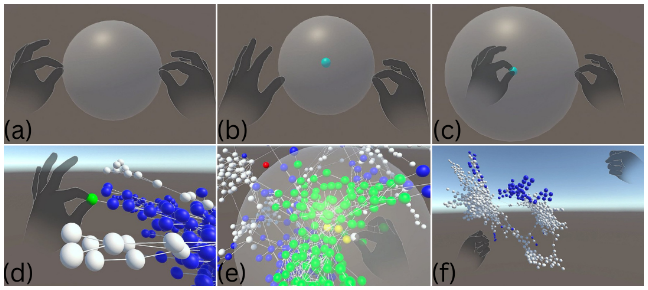

Figure 1: (a) sphere-SS; (b) sphere-low-AA; (c) sphere-high-AA; (d) single select; (e) single and sphere deselect; (f) rotate and scale.

GPU) ran the experiment application written in Unity version 2021.3.17f1. The VR app uses Oculus SDK to capture 3D positions from Quest 2 front cameras. All selections were accomplished by using pinch gestures. For single target selection or deselection, pinch gesture should have been inside the object. To create spheres, in SS both hands should have performed the pinch gesture at the same time. In AA conditions, the left hand creates the center of a sphere with a pinch gesture, and the right hand uses a pinch gesture to change the radius. When a sphere is created, a small red sphere appears at the top of that sphere; to delete that sphere and remove the selections inside that sphere, the red object should have been selected with the right-hand pinch. Bimanual fist gestures and movement while holding fists are used to rotate and scale Fig. 1.

Tasks Considering that multi-selection is a common class of actions in data exploration, we use multi-selection on a graph for our task. To ensure distinct groups for selection, we used a graph with non-overlapping groups and a ground truth partition. The graph data is the HCI subset of the Cora Citation Network [6]. To show the graph in 3D space in VR, we used a force-directed layout algorithm (Fruchterman-Reingold) to generate the 3D positions. The total number of nodes was 1053. The HCI citation graph included five groups Fig. 2. The task was to select as many members of the highlighted group as possible in 2 minutes while minimizing errors at the end of the trial. All tasks were performed in a standing position, allowing participants to walk around and select within-reach objects by either walking closer or using rotation and scaling interactions to position the objects nearby.

Design and Procedure The experiment employed a within-subjects design, with the primary factor being the technique, which had four levels: multi-single-select, sphere-SS, sphere-low-AA, and sphere-high-AA. Participants first completed a demographic form before performing multi-selection tasks using different bimanual interaction techniques. The order of bimanual interactions was determined by a balanced Latin square table. Each technique consisted of 2 blocks of 5 trials (groups) in random order. Besides selection, participants could rotate, scale, single deselect, and cross deselect (delete a previously created sphere) as illustrated in Fig. 1. During each trial, participants had 2 minutes to select the highlighted nodes, with the next trial starting immediately after 2 minutes, regardless of task completion. Participants were required to take at least a 5-second break after each trial but could rest longer if desired. At the end of the study, participants ranked the selection techniques according to their preferences.

The independent variables for this study are the types of selection techniques (conditions), with four levels: Sphere-SS, Sphere-low-AA, Sphere-high-AA, and Baseline (multi-single selection). We measured user performance using task success without error, efficiency (completion time for successful trials), and accuracy. Accuracy was defined as the total number of correctly selected nodes in a trial divided by the sum of the total number of incorrectly selected nodes and correctly selected nodes in a trial. In addition, we measured the F1 score and throughput. The F1 score was defined as F1 = 2·(P·R)/(P+R), where P (precision) is TP/(TP+FP) and R (recall) is TP/(TP+FN) [8]. Throughput was assessed in two ways: the number of correct selections divided by the completion time for successful trials without error, and the number of correct selections divided by the number of interactions (considering only selection interactions). We also analyzed the total number of interactions and the breakdown of five different types of interactions per trial.

3 Results

For all dependent variables, the distribution was not normal; so the replications of unique experimental conditions were represented by their median. A GLMM with standard repeated measures REML technique was used that handled participants as a random factor. For GLMM, the R lme4 package was used [1]. We report F-statistic using type III ANOVA with Satterthwaite approximation, and pairwise comparisons (using pooled variance) with Bonferroni correction. Initial level of significance (α) was set to 0.01. Friedman's ANOVA for within-subject comparisons, followed by post-hoc tests with Wilcoxon signed-rank tests for pairwise comparisons.

3.1 Testing H1

Success: A linear mixed effect model (LMM) found a significant main effect of METHOD, $F(1,771) = 65.944, p < .0001$, and GROUP, $F(4,771) = 128.499, p < .0001$ on SUCCESS. Significant interaction effects were also found for METHOD x GROUP, $F(4,771) = 10.593, p < .0001$. Pairwise comparisons found that *multi-single-select* resulted in less number of successful trials (N=94 out of 200) compared to *SS* (N=149), *low-AA* (N=128), and *high-AA* (N=140) Tab. 1. The total number of successful tasks across different conditions and groups is illustrated in Fig. 3.

Time: LMM found a significant main effect of METHOD, $F(1,263.27) = 60.420, p < .0001$, and GROUP, $F(4,263.94) = 157.975, p < .0001$ on COMPLETIONTIME. Significant interaction effects were also found for METHOD x GROUP, $F(2,263.11) = 17.742, p < .0001$. Pairwise comparisons found that *multi-single-select* resulted in significantly longer median completion times (78.42s) compared to *SS* (68.5s), *low-AA* (74.19s), and *high-AA* (72.89s) Tab. 1. The median completion times for successful trials across different conditions and groups are illustrated in Fig. 4.

Table 1: Pairwise comparisons of different dependent variables between multi-single-select and sphere methods (p-value<.0001).

Variable	Technique	Effect Size
Success	Multi-single, SS*	0.391
	Multi-single, Low-AA*	0.248
	Multi-single, High-AA*	0.335
Time	Multi-single, SS*	0.571
	Multi-single, Low-AA*	0.561
	Multi-single, High-AA*	0.614
Accuracy	Multi-single, SS*	0.858
	Multi-single, Low-AA*	0.867
	Multi-single, High-AA*	0.864
F1 score	Multi-single, SS*	0.861
	Multi-single, Low-AA*	0.855
	Multi-single, High-AA*	0.861
TP based on time	Multi-single, SS*	0.621
	Multi-single, Low-AA*	0.621
	Multi-single, High-AA*	0.604
TP based on time	Multi-single, SS*	0.868
	Multi-single, Low-AA*	0.868
	Multi-single, High-AA*	0.868

Accuracy: LMM found a significant main effect of METHOD, $F(1,371) = 144.425, p < .0001$, and GROUP, $F(4,371) = 9.981, p < .0001$ on ACCURACY. Significant interaction effects

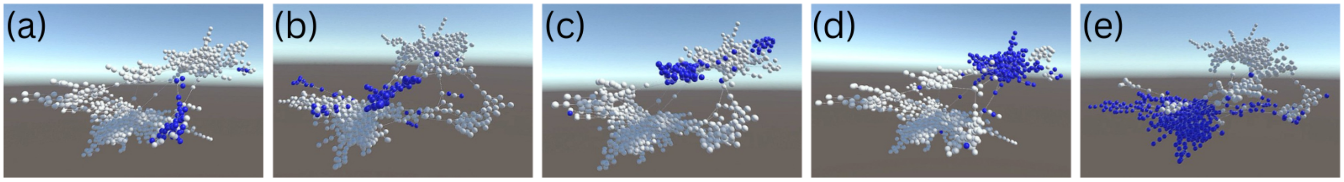

Figure 2: Groups: (a) wearable computers (39); (b) cooperative (79); (c) interface design (91); (d) multimedia (269); (e) graphics and VR (575).

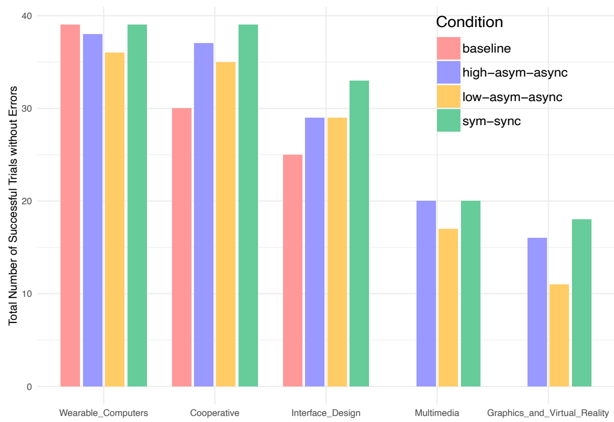

Figure 3: Successful trials for each condition and group.

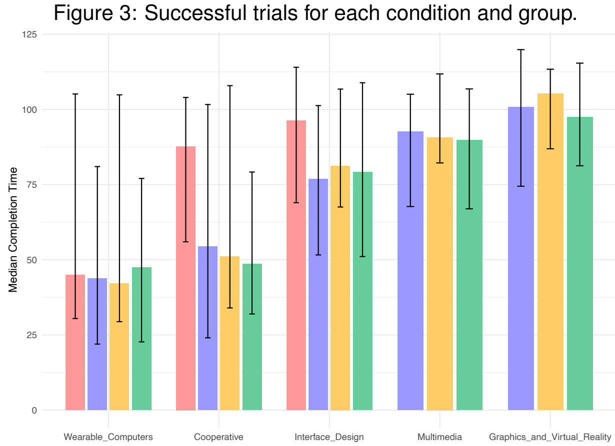

Figure 4: Completion Time of successful trials for each condition and group. Error bars = 95% CI.

were also found for METHOD X GROUP, $F(4,371) = 7.366, p < .0001$. Pairwise comparisons found that *multi-single-select* resulted in significantly higher accuracy (0.99) compared to *SS* (0.9), *low-AA* (0.85), and *high-AA* (0.86) Tab. 1.

F1 Score: LMM found a significant main effect of METHOD, $F(1,371) = 135.614, p < .0001$, and GROUP, $F(4,371) = 11.464, p < .0001$ on F1 SCORE. Significant interaction effects were also found for METHOD X GROUP, $F(4,371) = 3.395, p < .0001$. Pairwise comparisons found that *multi-single-select* resulted in significantly lower F1 Score (0.0.19) compared to *SS* (0.048), *low-AA* (0.0.49), and *high-AA* (0.052) Tab. 1.

Throughput: LMM found a significant main effect of METHOD, $F(1,264.21) = 27.568, p < .0001$, and GROUP, $F(4,265.41) = 909.699, p < .0001$ on THROUGHPUT BASED ON TIME. Significant interaction effects were also found for METHOD X GROUP, $F(2,263.92) = 9.532, p < .0001$. Pairwise comparisons found that *multi-single-select* resulted in significantly lower throughput based on time (0.9) compared to *SS* (1.6), *low-AA* (1.24), and *high-AA* (1.43) Tab. 1. LMM found a significant main effect of METHOD, $F(1,371) = 55.621, p < .0001$, and GROUP, $F(4,371) = 11.887, p < .0001$ on THROUGHPUT BASED ON SELECTION COUNT. Significant interaction effects were also found for METHOD X GROUP, $F(4,371) = 11.869, p < .0001$. Pairwise comparisons found that *multi-single-select* resulted in significantly lower throughput based on selection counts (0.8) compared to *SS* (4.13), *low-AA* (4.44), and *high-AA* (4.53) Tab. 1.

Summary The results support H1, indicating that sphere methods (both symmetric-synchronous and asymmetric-asynchronous) performed better than the baseline multi-single-select method in terms of success rate, completion time, and throughput. This suggests that bimanual techniques are more effective for multi-selection tasks in VR compared to serial selection methods.

3.2 Testing H2

Success: LMM found a significant main effect of GROUP, $F(4,566) = 63.196, p < .0001$ on SUCCESS. Pairwise comparisons revealed non-significant differences between number of successful trials of *SS*, *low-AA* and *high-AA*.

Time: LMM found a significant main effect of GROUP, $F(4,204.88) = 154.318, p < .0001$ on COMPLETIONTIME. Pairwise comparisons revealed non-significant differences between completion times of *SS*, *low-AA* and *high-AA*.

Accuracy: LMM found a significant main effect of GROUP, $F(4,266) = 26.488, p < .0001$ on ACCURACY. Pairwise comparisons found that *SS* resulted in significantly higher accuracy compared to *low-AA*, but revealed non-significant differences between accuracy of *SS*, and *low-AA*.

F1 Score: LMM found a significant main effect of GROUP, $F(4,266) = 10.74, p < .0001$ on F1 SCORE. Pairwise comparisons revealed non-significant differences between F1 Scores of *SS*, *low-AA* and *high-AA*.

Throughput: LMM found a significant main effect of GROUP, $F(4,260.5) = 806.617, p < .0001$ on THROUGHPUT BASED ON TIME. Pairwise comparisons found that *SS* resulted in significantly higher throughput based on Time compared to *low-AA*, but revealed non-significant differences between throughput based on Time of *SS*, and *low-AA*. LMM found a significant main effect of GROUP, $F(4,266) = 37.246, p < .0001$ on THROUGHPUT BASED ON SELECTION COUNT. Pairwise comparisons revealed non-significant differences between Throughput based on Selection Count of *SS*, *low-AA* and *high-AA*.

Summary The data did not show significant differences between symmetric-synchronous and asymmetric-asynchronous techniques regarding success rate, completion time, F1 score, and throughput based on selection count. Therefore, H2 is not supported, indicating that both bimanual techniques can be equally effective.

3.3 Testing H3

A series of LMMs analyzed various performance metrics across different groups. Significant main effects of GROUP were found on multiple metrics: SUCCESS with $F(4,371) = 43.991, p < .0001$, COMPLETIONTIME with $F(4,123.97) = 116.232, p < .0001$, ACCURACY with $F(4,171) = 18.551, p < .0001$, F1 SCORE with $F(4,171) = 5.522, p < .0001$, THROUGHPUT BASED ON TIME with $F(4,125.99) = 435.980, p < .0001$, and THROUGHPUT BASED ON SELECTION COUNT with $F(4,171) = 22.022, p < .0001$. Across all these metrics, pairwise comparisons revealed non-significant differences between the *low-AA* and *high-AA* groups, indicating similar performance levels in these conditions.

Summary The comparison between low and high tension asymmetric-asynchronous techniques showed no significant differences in performance metrics. Thus, H3 is not supported, suggesting that the level of kinesthetic tension does not significantly impact the effectiveness of these techniques in the context of this study.

3.4 Number of Interactions and Subjective Rankings

Fig. 5 illustrates the median interaction counts for various user interactions across multiple groups and conditions.

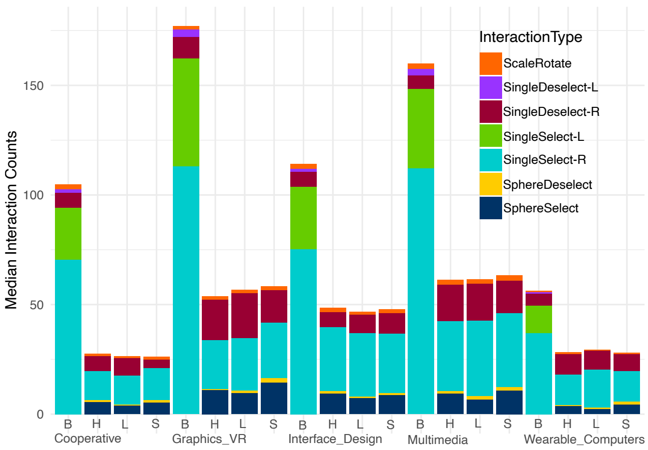

Figure 5: Median interaction counts for all Groups and Conditions. B: baseline (multi-single-select), H: high-AA, L: low-AA, S: SS

The counting of each condition in different preference positions revealed that *SS* was the most preferred condition, followed by *high-AA*,*low-AA*, and *multi-single-select*. A Friedman test indicated significant differences across these conditions overall ($\chi^2(3) = 40.56, p < 0.001$). Further pairwise comparisons using the Nemenyi-Wilcoxon-Wilcox test showed that significant differences were observed between *multi-single-select* and all other conditions. However, no significant differences were found between others. Fig. 6 displays the distribution of preferences for different techniques.

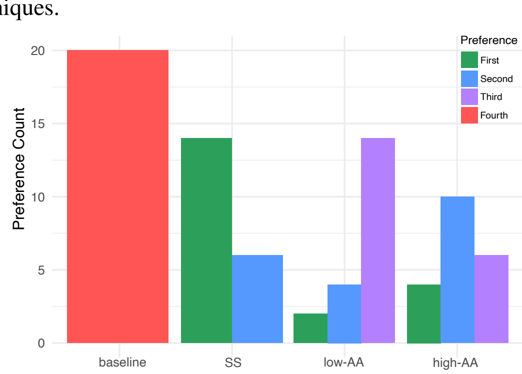

Figure 6: Preference counts based on subjective rankings.

4 Discussion

This study explored the effectiveness of bimanual, mid-air multi-selection techniques in VR. The results highlight insights into the performance and user preferences regarding these techniques.

Symmetric-Synchronous vs. Asymmetric-Asynchronous: The study revealed no significant differences in efficiency, accuracy, or throughput between symmetric-synchronous and high-tension asymmetric-asynchronous techniques for within-reach targets. This suggests that both techniques can be equally effective under certain conditions, possibly due to the intuitive nature of symmetric-synchronous actions and the refined control offered by asymmetric-asynchronous methods.

Preference for Symmetric-Synchronous Technique: Despite similar performance metrics, participants showed a marked preference for the symmetric-synchronous technique. This preference could be attributed to its natural alignment with human motor coordination, where simultaneous actions by both hands can lead to a more fluid and immersive experience in VR.

Challenges with Multi-Single Selection: The multi-single-select method was notably less effective across various metrics, including success rate and throughput. This underscores the limitations of serial selection in complex VR environments, where rapid and multiple selections are required to manipulate data efficiently.

Geometry and Mid-Air Multi-Selection Techniques in VR: In the context of data exploration in VR, the geometric configuration of data influences the efficiency of interaction techniques, particularly multi-selection methods [10]. We aimed to select a dataset for our study that encompasses these types, as shown in Fig. 2. Given that each geometry type presents unique challenges, they significantly impact the selection process, necessitating further investigation to fully understand their effects on mid-air interaction.

Future research should aim to extend the scope of this study by exploring the effectiveness of bimanual mid-air selection techniques in more diverse VR environments.

References

[1] D. Bates, M. Maechler, B. Bolker, S. Walker, R. H. B. Christensen, H. Singmann, B. Dai, and G. Grothendieck. Package 'lme4'. *Convergence*, 12(1), 2015. 2

[2] W. Buxton and B. Myers. A study in two-handed input. *ACM SIGCHI Bulletin*, 17(4):321–326, 1986. 1

[3] N. Chaconas and T. Höllerer. An evaluation of bimanual gestures on the microsoft hololens. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 1–8. IEEE, 2018. 1

[4] Y. Guiard. Asymmetric division of labor in human skilled bimanual action: The kinematic chain as a model. *Journal of motor behavior*, 19(4):486–517, 1987. 1

[5] A. Leganchuk, S. Zhai, and W. Buxton. Manual and cognitive benefits of two-handed input: an experimental study. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 5(4):326–359, 1998. 1

[6] A. K. McCallum, K. Nigam, J. Rennie, and K. Seymore. Automating the construction of internet portals with machine learning. *Information Retrieval*, 3:127–163, 2000. 2

[7] H. McLoone, K. Hinckley, and E. Cutrell. Bimanual interaction on the microsoft office keyboard. In *INTERACT'03*, pp. 49–56, 2003. 1

[8] R. A. Montano-Murillo, C. Nguyen, R. H. Kazi, S. Subramanian, S. DiVerdi, and D. Martinez-Plasencia. Slicing-volume: Hybrid 3d/2d multi-target selection technique for dense virtual environments. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 53–62. IEEE, 2020. 1, 2

[9] E. Olshannikova, A. Ometov, Y. Koucheryavy, and T. Olsson. Visualizing big data with augmented and virtual reality: challenges and research agenda. *Journal of Big Data*, 2:1–27, 2015. 1

[10] R. Stenholt. Efficient selection of multiple objects on a large scale. In *Proceedings of the 18th ACM symposium on Virtual reality software and technology*, pp. 105–112, 2012. 1, 4

[11] W. Stuerzlinger and G. Smith. Efficient manipulation of object groups in virtual environments. In *Proceedings IEEE Virtual Reality 2002*, pp. 251–258. IEEE, 2002. 1

[12] H. Surale. Barehand mode switching in touch and mid-air interfaces. 2020. 1

[13] A. C. Ulinski, Z. Wartell, P. Goolkasian, E. A. Suma, and L. F. Hodges. Selection performance based on classes of bimanual actions. In *2009 IEEE Symposium on 3D User Interfaces*, pp. 51–58. IEEE, 2009. 1

[14] A. M. Webb, M. Pahud, K. Hinckley, and B. Buxton. Wearables as context for guiard-abiding bimanual touch. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, pp. 287–300, 2016. 1