

EVALUATION OF MULTI-TOUCH TECHNIQUES FOR PHYSICALLY SIMULATED VIRTUAL OBJECT MANIPULATIONS IN 3D SPACE

Paulo G. de Barros¹, Robert J. Rolleston², Robert W. Lindeman¹

¹ Worcester Polytechnic Institute
100 Institute Road, Worcester, MA, USA.
{pgb, gogo}@wpi.edu

² Xerox Research Center
800 Phillips Rd, Webster, NY, USA.
rob.rolleston@xerox.com

ABSTRACT

Manipulation of objects in 3D space is not always a trivial task. When dealing with touchscreens, finding the correct set of mappings and interactions in three dimensions becomes even more of an issue. Current techniques evaluate 3D interaction regardless of physical properties of our world, aiming for task optimization instead of realism. The purpose of this paper is to evaluate three different types of 3D object manipulation techniques using multi-touch surfaces when the environment performs with a realistic physics behavior. The task used in the interface evaluation consists of organized objects in a box in a pre-specified way. Our results show that users tend to leverage from the physical properties of the environment to help them manipulate objects. Hence, the interface design that best approximates real-world manipulation performed best.

KEY WORDS

Virtual and augmented reality, user interface development, 3D selection and manipulation, multi-touch touchscreen interaction.

1. INTRODUCTION

The study of touch screen interaction for collaboration in tabletop and office environments has been thoroughly explored in Virtual Reality (VR). The research concepts of multi-hand, multi-person, multi-finger, multi-touch and multi-point are now commonplace [19]. Schönning *et al.* [19] summarizes the history of multi-touch screen and at which point we are now. It also poses questions that could or should be the topic of research in this area.

Despite the increasing number of touchscreen research publications, large touchscreens are not yet being widely used in daily tasks. Benko *et al.* [7] presents results of a survey performed with users of tabletop devices, mostly researchers in academia and industry. Their results indicate that touchscreen surfaces are often not used as intended. This discrepancy is mostly because of ergonomic issues, text entry difficulties, and imprecision in finger-pointing.

When compared to tangible interfaces for object acquisition and manipulation [20], touchscreen interfaces had inferior interaction performance because they were less natural to use. In addition, error in interactions with multi-touch interfaces tends to occur because of the dislocation of the finger's centroid on screen during finger release.

Another comparison between mouse input and multi-touch input using one or two points has shown that, in a

docking task and with only one point on screen, mouse has more accuracy, while touch is slightly faster [9]. For two points, touch is better because controlling two mouse pointers on screen slows down mouse interaction. However, fatigue and collaboration must be considered when choosing either input mode for a certain application.

Based on these examples, and as for most interaction devices, multi-touch interfaces seem to have their pros and cons. Their suitability for a certain application will depend on the application requirements for manipulation accuracy, simultaneous input and naturalness among others.

The current work aims at evaluating three types of multi-touch input interfaces for 3D object manipulation in an environment with realistic physics behavior. The task of packing objects in a box was used for interface evaluation. This task consists in assembling collections of objects such as books, notepads, CDs and pens in a single packaged unit. The three interfaces were designed, built, tested and compared in terms of performance and subject preference.

This paper is divided as follows. Section 2 highlights the related work. Section 3 describes the three interface designs. Section 4 delineates the interface evaluation methodology. Section 5 reports the user study results. Section 6 makes an analysis of such results. Last, section 7 draws conclusions by identifying key points in our results.

2. RELATED WORK

A large body of research already exists in the area of multi-touch interface interaction. Guidelines and various interface designs have been explored for 3D user interaction and multi-user tabletop collaboration [21]. Apted *et al.* [3] reviews heuristics for tabletop applications and proposes their own set of heuristics for interaction with tabletops.

We *et al.* [26] presents a simple study of properties in multi-touch interaction that should be common to applications. Three features that can be used and combined to generate interesting interaction are highlighted:

- *Registration*: when gestures are identified in the device;
- *Relaxation*: once a gesture is identified, the user does not need to maintain the gesture during the entire interaction process.
- *Reuse*: one gesture can represent different interactions depending on the state of the application. This reduces complexity of the interface and potentially its cognitive load if the shifts in gesture representation are intuitive.

Witzel [23] has proposed and tested graph schemes for identifying multimodal interactions using single-touch devices, cyclic menus and arrows in VR applications.

PDA's and smart phone interfaces have also been the subject of touchscreen interface research due to their limitations in size, resolution and applicability [1].

Touch tables and walls have been explored in ethnographic studies, interface design and development and user testing in architecture [25], astrophysics [21][10], image manipulation, [26] among others.

To simplify interface use and speed up its learning, exemplified hand gestures on screen, videos and help menus and other standard systems have been used [22].

Touchscreens' bidimensional input has been used in the improvement of current interface functionalities such as scrolling [4], selection [8][2], panning, zooming [14] and menu access [5] and 3D navigation [10].

Because of this bidimensional interaction nature of screens, using them for 3D object manipulation involves remapping controls through multi-touch, gestures, input detection techniques or interaction modes [12]. But, optimizing the match of input and actions is not trivial.

A thorough exploration of the intuitiveness and simplicity of each interaction from both HCI experts and user perspectives [16][17][18][24] is a common approach. Hancock *et al.*[11] claim that using manipulation with fewer degrees of freedom at a time is generally more intuitive and that using more fingers is generally better.

Martinet *et al.* [15] has evaluated two techniques for 3D object positioning using multi-touch screens in an docking task. Kruger *et al.* [13] has proposed a single-finger object rotation and translation interface.

Many 3D user interaction techniques used in other input interfaces could be remapped to multi-touch screens. Good candidates for this are the balloon selection map [6] and the Go-go navigation technique [16].

The interfaces proposed here used as a basis many of the current recommendations of the above authors in a way that could best facilitate virtual box packing. However, in our case, we evaluate the effect on user performance using such interfaces when physics simulation is added as a component of the 3D application.

3. INTERFACE DESIGNS

Three interface designs were created, each with different levels of multi-touch interactivity and selected from a set of six designs. The number of interfaces selected was reduced from six to three so that experimental sessions would be short enough to make a user study feasible.

These interface designs were tested and optimized based on informal feedback during a short pilot study. They leverage from multi-touch screen interaction in different degrees. The gizmo interface requires only one finger for operation, similar to a mouse interaction. This interface was used as our control case. The Uni-manual Constrained Control (UCC) one allows manipulation using two fingers, one in each hand. It is equivalent to having to mouse pointers. The Move&Turn technique uses three fingers: two in one and one in the other hand.

The physics simulation used was designed to simulate basic physics interactions such as collisions. Objects did not have momentum in the x-y plane due to user screen. However, an object falling under the force of gravity could collide with other objects and cause these objects to have a realistic physical response. Two of the three interfaces had an active gravity component, while one had the ability to turn off gravity and temporarily leave objects suspended in space.

3.1 Gizmo Interface

Most if not all 3D modeling tools nowadays leverage from a gizmo to manipulate objects. A similar interface was used in our study as the control group for measuring effectiveness of the other two interface designs proposed.

The gizmo interface, presented in Figure 1, consists of three colored bidirectional arrows perpendicular to each other, and each containing spheres close to their tips. They are attached to each other by their central point. Attached diagonally to them are two squares, one at a vertical orientation and the other oriented horizontally. Only one finger is necessary to use this interface.

Touching and dragging a finger over any of the colored arrow tips translates the selected object along the directions the arrows are pointing to. The red arrow uses a horizontal sliding gesture on screen for translation while the blue and green arrows use a vertical sliding gesture.

By touching and sliding the colored spheres, the user can rotate the selected object. Horizontally sliding the finger on the green and blue spheres and vertically on the red spheres rotates the object around the axle (arrow) the sphere is attached to. The direction of rotation would vary according to the direction of the finger slide.

Touching the horizontally aligned square labeled "XZ Trans.", and sliding the finger on the screen translates the selected object along the horizontal plane as if using both the red and green arrows simultaneously.

Similarly, touching the vertically aligned square labeled "XY Trans.", and sliding the finger on the screen translates the selected object vertically plane as if using both the red and blue arrows simultaneously.

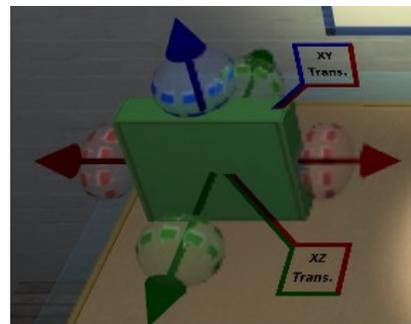


Figure 1. Gizmo interface with a green box selected.

The selected object can be kept up in the air by releasing all fingers off the screen. This allows the user to rest his hands while planning for the next move. A finger touch outside the gizmo or the selected object deselects the object and drops it over the nearest horizontal surface.

3.2 Move & Turn Interface

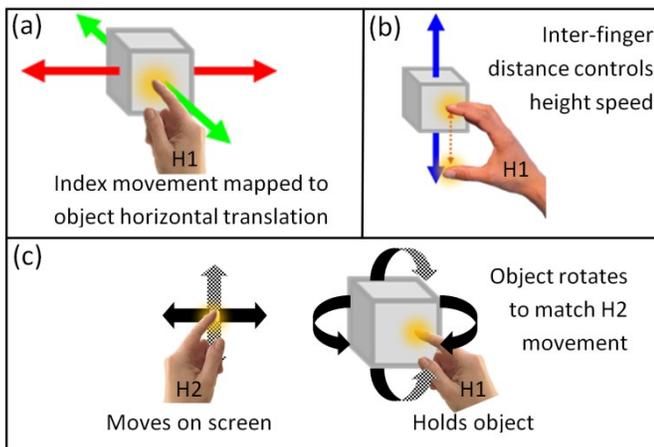
The Move & Turn interface distinguishes interaction modes by the amount of fingers on screen and their spatial relationship. At most three fingers are required: two fingers from one hand and one finger from the other. In our study, the fingers used were the index ($H1index$) and thumb ($H1thumb$) fingers from the dominant hand and the index finger from the non-dominant hand ($H2index$).

Touching an object with $H1index$ selects and holds it. Releasing that finger from screen releases and drops the object. Therefore, this interface does not allow an object to be kept up in the air without having to touch the screen and does not allow hand rest between manipulations. Sliding $H1index$ horizontally translates the object, equivalent to using “XZ Trans.” square of the Gizmo interface (see Figure 2a).

Height is controlled by varying the distance between $H1index$ and $H1thumb$ when they are both touching the screen. The variation of the distance between these two fingers is mapped to an increment or decrement in the upward moving speed of the selected object (Figure 2b). For example, by touching both of these fingers on screen, increasing the distance between them and then holding the fingers in that position, the selected object would move up at a constant speed. The object is stopped by releasing the $H1thumb$ or by returning the fingers position to a distance equivalent to the one when they first touched the screen.

The $H2index$ controls the rotation of a selected object that is being held by $H1index$ (Figure 2c). Sliding $H2index$ horizontally rotates the selected object around the vertical axle, equivalent to using the Gizmo interface blue spheres. Sliding it vertically rotates the object around the horizontal axle perpendicular to the camera direction, equivalent to using the Gizmo interface red spheres.

Figure 2. Move & turn interface (right-handed person):



(a) horizontal translation by sliding $H1index$ along screen surface; (b) changing height by varying height speed based on distance between $H1index$ and $H1thumb$; (c) rotating object using $H2index$ finger while holding object with $H1index$.

3.3 Uni-manual Constrained Control (UCC) Interface

The UCC interface defines three interaction modes, each with two degrees-of-freedom (DOF) at a time. The two DOF input is given by horizontally and vertically sliding the finger holding the object along the screen surface. Similar to the Move&Turn interface, releasing all fingers from the screen releases and drops the object. Hence, UCC does not allow holding the object in the air without touching the screen for relaxation between manipulations.

Two fingers are required to use this interface. The first finger selects and holds the object, similar to the Move & Turn interface. The second cyclically switches between interaction modes. Thus, this interface reuses gestures to perform different actions. In our study, the fingers used for selection and interaction mode switching were the dominant hand’s index ($H1index$) and thumb ($H1thumb$) fingers respectively. Interaction modes were identified by an icon placed on top of $H1index$ (Figure 3).

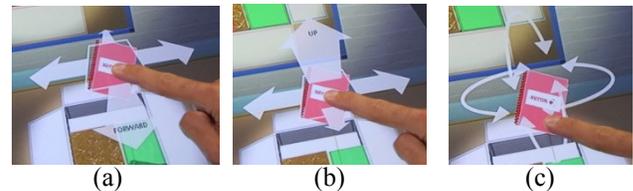


Figure 3. UCC interface (right-handed person): (a) horizontal translation mode; (b) vertical translation mode; (c) rotation mode.

The first interaction mode horizontally translates the selected object, equivalent to using the Gizmo interface “XZ Trans.” square. The second mode vertically translates the object along a plane equidistant to the camera, equivalent to using the Gizmo interface “XY Trans.” square. The third mode rotates the object around the vertical and horizontal axes, similar to the rotations performed with $H2index$ in the Move & Turn interface.

4. HYPOTHESIS

We claim that the closer an interface represents the task as if being performed in real-world, the better it should perform in a 3D environment with physics simulation. In our study, the task, described in section 5.2, is packing-related, where a set of objects must be organized in a box.

With that in mind, we hypothesize that the Move&Turn interface should outperform the other two interfaces (H1), since it most-closely and naturally approximates performing the task in the physical world.

In addition, we hypothesize that the UCC interface should have the second best performance (H2), since it more closely and naturally approximates performing the task in the physical world than the gizmo interface.

5. METHODOLOGY

In order to evaluate the performance of the interfaces for a packing task, a user study was carried out. A within-subjects experiment was run. The independent variable was the touchscreen interface used in performing the task.

5.1 Virtual Environment & Equipment

The virtual environment (VE) where the task was performed was a table with a white box and 6 objects around it as presented in Figure 4.

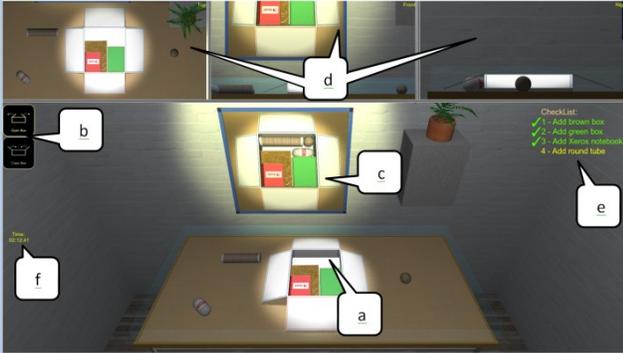


Figure 4. Virtual environment for user study: (a) white box where 6 objects should be placed. Three are already inside the box; (b) buttons for opening and closing white box; (c) picture showing how objects should be organized in the box; (d) Top, front and right perspectives from left to right; (e) Check list to hint the user on the order objects should be placed. (f) Timer measuring subjects' task time.

The VE was developed with the Unity3D game engine, version 2.6.1f3. It was responsible for scene management, rendering and physics simulation. All objects had a mass, collided with each other and were subject to gravity when not held by the subject's touch.

The machine used was a Dell Studio XPS with Core™ i7 processors (2.8GHz), 12 GB RAM with Windows® 7 Professional 64-Bit and a GeForce GTS 240 graphics card. The multi-touch device was a 46 inches HD TV with a PQLabs G² system installed on top of it.

5.2 Task

The experiment task was to open the white box in the middle of the table top, place the other 6 objects that are on the table inside this box and then close it. However, the objects had to be organized inside the box in a certain way according to what was shown by a picture on the wall in front of the table. Subjects were asked to complete the task as fast and accurately as they could.

A checklist showing the order with which objects should be placed in the box was presented on the screen.

Even though the camera perspective was fixed, in order to help subjects place and align objects, three smaller views with top, front and lateral camera perspectives of the table top were displayed on top of the screen. However, objects could not be manipulated through these views, but only through the main view.

5.3 Data Collection

Subjects' demographical data was collected as well as subjective impressions about the interface. Most subjects (30 out of 37) had their experiment recorded on video.

The independent variable was the type of interface used and the dependent variables were the time to finish the task and the object placement accuracy inside the box.

Each of the 6 objects had from 1 to 6 points across its surface whose positions were measured every time the experiment was over. At the end of each task with one of the interface, these points were compared to their optimal positioning, that is, when objects were placed exactly as in the picture presented on the wall. This error was recorded for each session, which consisted of a combination of a subject and an interface type. This error determined the positioning accuracy for each of the object's that were placed in the box by the user.

Other data was also collected. The periods of time spent in the training room and performing individual types of rotations and translations during task were also recorded. Because of the nature of each interface, the variety of the collected information changed (Table 1).

Table 1: Time-related information collected for interfaces.

Time Spent on Performing:	Gizmo	Move&Turn	UCC
Training	X	X	X
Any translation	X	X	X
Any rotation	X	X	X
Translations on X	X		
Translations on Y	X	X	
Translations on Z	X		
Translations on XY plane	X		X
Translations on XZ plane	X	X	X
Rotations around X	X		
Rotations around Y	X		
Rotations around Z	X		

5.4 Experimental Procedure

When subjects arrived in the lab, they were requested to carefully read the experiment instructions and ask questions about the task to be performed. They were allowed to stop participating in the experiment at any time whenever they desired for any reason.

Next, a pre-questionnaire was applied to collect information about age, gender, experience with videogames or 3D applications and experience with touch screens or touchscreen devices such as smart phones. The possible answers for the last two questions were daily, weekly, seldom or never.

The subject was then put in a training room to learn how to operate the first interface out of the three that were designed. The task in the training room was to place a single object inside the box. Whenever the subject fully understood how to operate the interface, the subject was moved to the experiment VE where he had to perform the task of adding the 6 objects in the box using the interface previously presented and learned during the training session. Once the task was completed for this interface, the subject was allowed to take a 5-minute break.

After the task with the first interface was over, the same procedure of training room practice, performing task and taking a break was done for the other two interfaces.

The exact same task was requested to be performed for the three interfaces. To avoid bias caused by accumulated learning about the task, how to use the touch-screen or the interaction interfaces during the

experiment, the order in which interfaces were presented for subjects was randomized using a Latin Square.

Once a subject's sessions were over, he was asked to fill-in a post-questionnaire expressing their impressions about the interfaces and the experiment (see Figure 5).

1. Sort the interfaces according to your preference, the first being the best one for performing the assigned task.
2. Sort the interfaces according to your preference, the first being the best one for performing the assigned and other 3D manipulation tasks or activities. Please list other tasks or activities that led you to this choice.
3. Please provide any comments about the Move & Turn (Multi-hand) interface.
4. Please provide any comments about the UCC interface.
5. Please provide any comments about the Gizmo interface.
6. Do you have any comments about the experiment in general?

Figure 5. Questions presented in post-questionnaire.

No questions were allowed during the task or the breaks. Subject's questions about the experiment and its purpose were answered only when the experiment for that subject was finished.

6. RESULTS

A total of 37 subjects (12 women and 25 men) volunteered for the experiment. Their age ranged from 19 to 63. Six subjects were left-handed.

A careful analysis using Analysis-of-Variance (ANOVA, $p=0.05$) of the population demographics information collected has shown that age, gender, handedness, videogame and 3D application experience and touchscreen experience had neither an effect on the time to perform the task nor in the accuracy with which the objects were positioned in the box.

During the experiments some subjects had problems with the physical simulation and that led to a slight increase in their total time. An analysis of the experimental data without data from these subjects still led to the exact same results obtained from an analysis of the data with all subjects. This indicates that such problem did not impact results in a statistically significant manner. Hence, all users were kept in our final analysis results.

Comparisons of results amongst the three interfaces were done pair-wisely using ANOVA ($p = 0.05$).

The results for object positioning have shown no statistically significant difference in accuracy among the three interfaces. The presumed cause for this was the fact that subjects used the physical properties of the environment to correctly align a selected object inside the box by having it collide with objects around it. This behavior caused the physics simulation engine, not the change in interface modalities, to be the factor that mostly contributed to determining the accuracy with which objects were placed in the box.

The results for task time are presented in Figure 6. A statistically significant difference was detected between times for pairs of interfaces: (Move&Turn, UCC; $F = 7.42, p = 8.08 \times 10^{-3}$) and (Move&Turn, Gizmo; $F = 12.63, p = 6.74 \times 10^{-4}$). This seems to indicate that Move&Turn

allowed subjects to perform the task faster. This result supports our first hypothesis (H1).

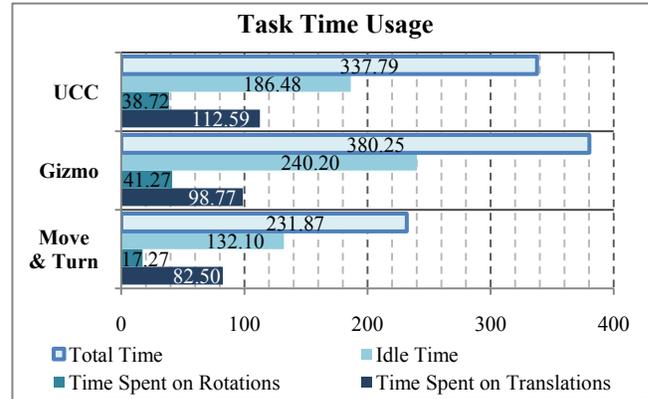


Figure 6. Average task time in seconds for each interface.

Figure 6 also shows the time spent doing translation and rotations for each interface. When using the Move & Turn interface, subjects had a statistically significant decrease in rotation time compared to the other two interfaces (Move&Turn and UCC: $F = 25.18, p = 3.63 \times 10^{-6}$; Move&Turn and Gizmo: $F = 15.76, p = 1.68 \times 10^{-4}$). This might be due to the decoupling between translations and rotations between hands in the Move&Turn interface. But, it might also be due to the closest relationship between the interactions in this mode and the actions that were to be performed to accomplish the task in the real world. When using the UCC interface, subjects had a statistically significant increase in translation time compared to the Move&Turn interface ($F = 6.59, p = 0.01$). The alleged cause for this was the constant confusion in interaction modes reported by subjects.

For the Move & Turn interface, 84% of the translation time was spent in changing height, compared to translating objects along the table. This is a strong indication that changing height with this interface was a non-trivial task while organizing the objects when they were inside the box wasn't.

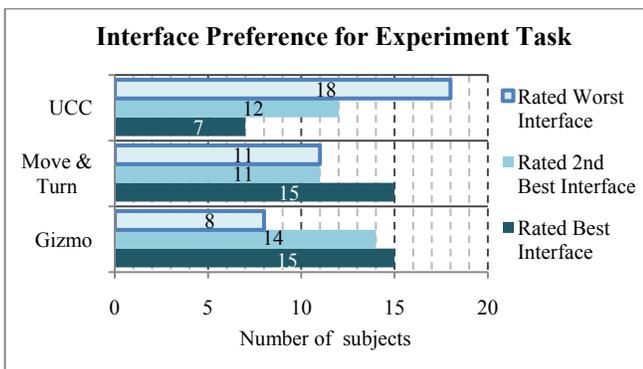
For the Gizmo interface, most of the translation time was fairly distributed amongst the three axes, the red axle having the most time dedicated to it. In terms of rotation time, 60% of the time was spent on vertical rotations for the Gizmo interface. These variations are claimed to have been caused by the nature of the task and how objects were organized on the scene. Notice that more than twice the amount of time was spent in rotations for this interface compared to what was spent when using the Move&Turn interface. This might be because the high abstraction level of the Gizmo interface does not incite the user to leverage from the physical properties available in the VE. Such properties could make object alignment tasks easier.

For the UCC interface, translation time was equally subdivided between vertical and horizontal translations.

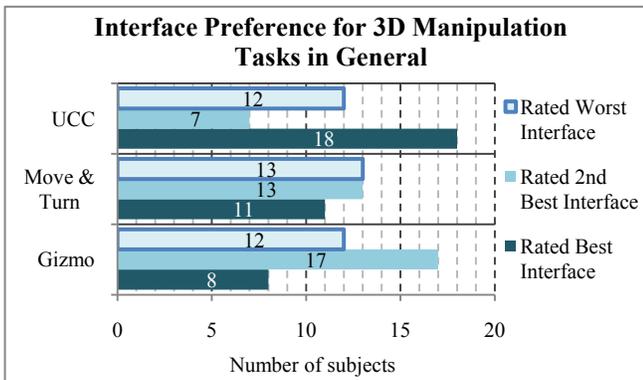
An effect of the order with which interfaces were presented on the users' improvement in terms of task time was only detected for the UCC interface. Subjects that were presented with UCC as their first interface spent a

statistically significant larger amount of time performing the task than subjects who were presented with UCC as their second or third interface.

The subjective results of the post-questionnaire have shown a tie of preference between the Move&Turn and the Gizmo interface for performing the experiment task (answers from question 1 in Figure 5). Interestingly, when asked the same question but for any general manipulation task (answers from question 2 in Figure 5), most votes went for the UCC interface followed by the gizmo interface. Both these results are presented in Figure 7a and 7b respectively. According to subjects, some of the reasons that affected their decision in changing the interface preference order were envisioning the use of the interfaces in 3D modeling applications, videogames and smart phones.



(a)



(b)

Figure 7. Subjects interface preference for (a) the experiment task and (b) other tasks in general.

Table 2 lists the most mentioned positive and negative comments for each of the interfaces. Each of them was mentioned by at least 3 subjects. Positive comments are presented in *italics>. General compliments were also done for all the interfaces, but the gizmo interface had the highest amount. These comments were not included here.*

The gizmo interface seemed to be recognized as the easiest to learn and use, but also the interface that allows more precise manipulation of objects. On the other hand, the Move & Turn interface seemed to be recognized as the fastest one to perform the task.

Table 2: Common subject comments for interface designs.

Move & Turn:	Gizmo:
<ul style="list-style-type: none"> Height manipulation is hard to learn or counter-intuitive Rotation is hard to learn or counter-intuitive <i>Fast to use</i> Dislike having to hold objects with finger to avoid them falling or finds it tiring 	<ul style="list-style-type: none"> Arrows overlapping is confusing <i>Likes not having to hold objects with finger.</i> <i>Rotation was easy</i> <i>Good for precision work</i> <i>Simple, easy to use</i> Gizmo is large and blocks view Had problems figuring out which part of the gizmo to touch to perform a certain rotation <i>Gives more insight on the effects of the manipulation</i> Had problems figuring out which part of the gizmo to touch to perform a certain translation
UCC:	
<ul style="list-style-type: none"> Had problems with accidentally switching modes Rotation is hard to learn or counter-intuitive Expected modes order to reset for a new object Time consuming, slow to use. bulky, annoying, frustrating 	

A comparison for each subject between responses for preferred interfaces and interface with which he/she performed better led to the results in Figure 8. It indicates that there was higher compatibility between subjects opinions and subjects performance for the Move& Turn interface, followed by the gizmo interface.

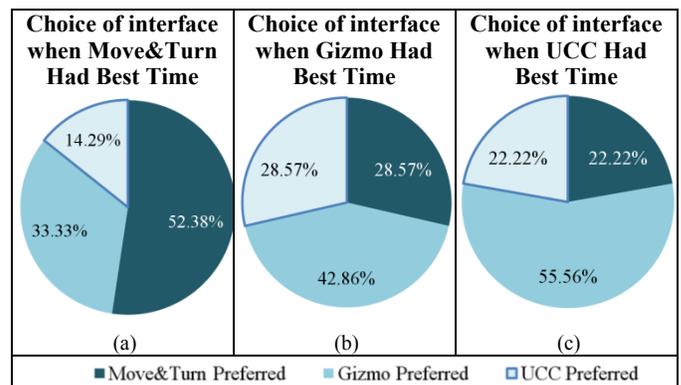


Figure 8. Subjects' choice of interface when subjects performed best with (a) Move & Turn, (b) Gizmo and (c) UCC interfaces.

7. CONCLUSION

The result from the user study has confirmed the first hypothesis that the use of the Move & Turn interface leads to an improvement in performance for the packing task in terms of time in an environment where physics simulation exists.

Nevertheless, the hypothesis was inconclusive with respect to performance improvement related to positioning of objects inside the box. We believe that the cause was the prevalence of the physics simulation control of the objects position over the interaction interfaces control.

Our second hypothesis was rejected. The UCC interface was not only the one with poorest performance but also the least preferred for performing the packing task. It is believed that the problems of switching

interaction modes by accidentally touching the screen and the fact that the interaction mode state was persistent at every new interaction, instead of being reset, led to the poor performance of this interface.

The Move & Turn interface has proved to be the best interface for performing the packing task according to our results. However, when compared to the Gizmo interface, a longer period of training seems to be necessary so that subjects can get habituated to it. In addition, the interface should and will go through important modifications and evaluations in the next iterations, such as the effect of changing height manipulation from rate to absolute control and making the rotation behavior more natural.

Having the rotation being kept separate from the hand holding the object for the Move & Turn interface seemed to have reduced user's cognitive load on the user as the rotation times seem to indicate.

More importantly, when analyzing the video of subjects using this interface, it was noticed that subjects had a tendency to use the physical properties of the environment to their advantage when aligning objects inside the box. They would perform small bumps and collisions with the objects in order to align the object being held with the box it is being put in. *This reproduces a more natural behavior that is done in routine packaging tasks in the real world.* Although, this behavior was also slightly noticed with the UCC interface, it was not noticed as evidently with the gizmo interface. Again, we believe that the cause for such a difference was the fact that the gizmo interface graphically represents more closely the abstract geometrical concepts of rotation and translation and when using this interface the fact that objects floated in the air when selected. Both of these factors may have led to a distancing from the potential physical manipulation realism embedded in the virtual scene.

On the other hand, the gizmo interface allowed translation and rotation of the selected object one axle at a time. This gave the subjects the power to more easily and precisely plan and implement their manipulations, though in a slower fashion compared to the Move & Turn interface.

Moreover, the gizmo interface was graphically self-explanatory whereas the manipulation for the Move & Turn interface had to be already known by the user or recognized by trial and error in real-time. Yet, the gizmo interface cluttered the screen area around the selected object. Despite being semi-transparent, it had made visibility difficult during alignment tasks. It also blocked selection of other small objects near the currently selected object.

In summary, it appears that our result seem to go in accordance with what was claimed by Hancock *et al.*[11] and referred in section 2 of this paper. We can claim that fewer degrees of freedom such as in the gizmo interface was preferred by users, but that using more fingers have indeed improved performance as was the case of the Move&Turn interface.

The gizmo interface seems to have a higher level of abstraction and is slow for the packaging task with physics simulation, but it is easy to learn and use. It also seems to be more easily extensible to other tasks as it has been used by many 3D modeling applications. It should probably be the best choice for a manipulation tool with a variety of different task environments when physics realism is not present or necessary.

The Move & Turn interface closely represents the actions common to a packing task and, by pushing the subject to use natural behaviors to perform the tasks, leads to an overall better performance when physics simulation is involved. However, it takes longer to learn because of the lack of hints on its interactivity and because of the necessary effort in remapping the depth dimension to the 2D domain of the multi-touch surface. In addition, there is no evidence that the Move & Turn could potentially be used with physics simulation in tasks other than the one here studied.

Last, the UCC interface seems to be a good balance between the other two interfaces for more general manipulation as most subjects have suggested, but its mode toggling needs to be improved and validated in further experiments.

8. FUTURE WORK

This paper has presented a first experiment in using multi-touch devices with a realistic physics simulated environment to measure the impact of the use of physics in interface design for the task of packing a box with an assortment of items. It contained important results that will guide the group towards the creation of new and improved multi-touch interfaces for virtual environments where physics laws apply.

We also plan on adding in a more accurate physics model to the interface by adding momentum to the objects as they are released by the user. This would require assigning speed to the objects based on the gesture speed of the user interaction.

A longer-term user study with updated interfaces needs to be carried out so that the interfaces can be evaluated by performing tasks with experienced users. This will allow us to collect data about these interfaces when performing this task and a variety of other more general ones in a way that will not be affected by subject levels of practice due to the differences between the interfaces learning curves.

A comparison of these interfaces for performing the same task with and without physics simulation also needs to be performed. It will allow the detection of the variation in user preference because of the physics simulation itself.

ACKNOWLEDGEMENTS

We would like to thank Xerox Corporation for providing the opportunity that made possible performing this research. In particular we thank the other project members for providing feedback on the interface designs. We would also like to thank all subjects who volunteered for

the user study reported here. Last, we would like to thank Worcester Polytechnic Institute for granting permission for performing this research.

REFERENCES

- [1] P. M. Aoki, A. Hurst & A. Woodruff, Tap Tips: Lightweight discovery of touchscreen targets, *Extended Abstracts of the ACM SIGCHI Conf. on Human Factors in Computing Systems (CHI)*, Seattle, WA, 2001, 237-238.
- [2] P.-A. Albinsson & S. Zhai, High precision touch screen interaction, *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems*, Ft. Lauderdale, Florida, USA, 2003, 105 – 112.
- [3] T. Apted, A. Collins & J. Kay, Heuristics to support design of new software for interaction at tabletops, *CHI Workshop on Multitouch and Surface Computing*, 2009.
- [4] K.W. Arthur, N. Matic & P. Ausbeck, Evaluating touch gestures for scrolling on notebook computers, *Extended Abstracts on Human Factors in Computing Systems (CHI)*, Florence, Italy, April 05-10, 2008.
- [5] G. Bailly, E. Lecolinet & Y. Guiard, Finger-count & radial-stroke shortcuts: 2 techniques for augmenting linear menus on multi-touch surfaces, *Proc.s of CHI*, 2010, 591-594.
- [6] H. Benko & S. Feiner, Balloon Selection: A multi-finger technique for accurate low-fatigue 3D selections, *Proc. of IEEE Symp. on 3D User Interfaces*, Charlotte, North Carolina, USA, March 10–11, 2007, 79–86.
- [7] H. Benko, M.R. Morris, A.J.B. Brush & A.D. Wilson, Insights on interactive tabletops: a survey of researchers and developers, *Technical Report*, Microsoft Research Center, number: MSR-TR-2009-22, 2009.
- [8] H. Benko, A.D. Wilson & P. Baudisch, Precise selection techniques for multi-touch screens, *ACM CHI Conf. on Human Factors in Computer Systems*, 2006, 1263-1272.
- [9] C. Forlines, D. Wigdor, C. Shen & R. Balakrishnan, Direct-touch vs. mouse input for tabletop displays, *Proc. of CHI*, 2007, 647–656.
- [10] C.W. Fu, W.B. Goh & J.A. Ng., Multi-touch techniques for exploring large-scale 3D astrophysical simulations, *Proc. of the 28th Int'l Conf. on Human Factors in Computing Systems (CHI)*, 2010, 2213-2222.
- [11] M. Hancock, S. Carpendale & A. Cockburn, Shallow-depth 3D interaction: Design and evaluation of one-, two- and three-touch techniques, *Proc. of CHI*, 2010, 1147–1156.
- [12] O. Hilliges, Bringing the physical to the digital: a new model for tabletop interaction, *Dissertation*, LMU München: Faculty of Mathematics, Computer Science and Statistics, 2009.
- [13] R. Kruger, S.T. Carpendale, S.D. Scott & A. Tang, Fluid integration of rotation and translation, *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems (CHI)*, 2005, 601-610.
- [14] S. Malacria, E. Lecolinet & Y. Guiard, Clutch-free panning and integrated pan-zoom control on touch-sensitive surfaces: the cyclostar approach, *Proc. of CHI*, 2010, 2615-2624.
- [15] A. Martinet, G. Casiez & L. Grisoni, Design and evaluation of 3D positioning techniques for multi-touch displays, *Research Report*, RR-7015, INRIA, 2009.
- [16] I. Poupyrev, M. Billinghurst, S. Weghorst & T. Ichikawa, Go-Go interaction technique: Non-linear mapping for direct manipulation in VR, *Proc. of UIST*, 1996, 79-80.
- [17] J. L. Reisman, P. L. Davidson & J. Y. Han, A screen-space formulation for 2D and 3D direct manipulation, *Proc. of the Annual ACM Symp. on User Interface Software and Technology (UIST)*, 2009, 69-78.
- [18] M.R. Morris, J.O. Wobbrock & A.D. Wilson, Understanding users' preferences for surface gestures, *Proc. of Graphics Interface (GI)*, Ottawa, Ontario, Canada, May 31-June 2, 2010, 261-268.
- [19] J. Schöning, A. Krüger & P. Olivier, Multi-touch is dead, long live multi-touch, *CHI Workshop on Multitouch and Surface Computing*, Boston, MA, USA, April 4-9, 2009.
- [20] P. Tuddenham, D. Kirk & S. Izadi, Graspables revisited: Multi-touch vs. tangible input for tabletop displays in acquisition and manipulation tasks, *Proc. of ACM Conf. on Human Factors in Computing Systems (CHI)*, 2010, 2223-2232.
- [21] D. Wigdor, H. Jiang, C. Forlines, M. Borkin & C. Shen, WeSpace: the design development and deployment of a walk-up and share multi-surface visual collaboration system, *Proc. of CHI*, 2009, 1237-1246.
- [22] D. Vanacken, K. Luyten & K. Coninx, TouchGhosts: Guides for improving visibility of multi-touch interaction, *Workshop on Multitouch and Surface Computing (CHI)*, Boston, USA, April 4, 2009.
- [23] M. Witzel, Graph-based multimodal interaction for design review in virtual and augmented reality, *PhD dissertation*, University of Trento, Italy, December 2007.
- [24] J.O. Wobbrock, M.R. Morris & A.D. Wilson, User-defined gestures for surface computing, *Proc. of the 27th Int'l Conf. on Human Factors in Computing Systems (CHI)*, Boston, MA, USA, 2009, 1083-1092.
- [25] M. Wu & R. Balakrishnan, Multi-finger and whole hand gestural interaction techniques for multi-user tabletop displays, *ACM Symp. on User Interface Software and Technology (UIST)*, 2003, 193-202.
- [26] M. Wu, C. Shen, K. Ryall, C. Forlines & R. Balakrishnan, Gesture registration, relaxation, and reuse for multi-point direct-touch surfaces, *IEEE Int'l Workshop on Horizontal Interactive Human-Computer Systems (TableTop)*, 2006, 185-192.