Multi-touch RST in 2D and 3D Spaces: Studying the Impact of Directness on User Performance

Sebastian Knoedel, Martin Hachet

Multi-touch RST in 2D and 3D Spaces: Studying the Impact of Directness on User Performance

Sebastian Knoedel* Martin Hachet†

INRIA Bordeaux - Université de Bordeaux CNRS (LaBRI)

ABSTRACT

The RST multi-touch technique allows one to simultaneously control *Rotations*, *Scaling*, and *Translations* from multi-touch gestures. We conducted a user study to better understand the impact of directness on user performance for a RST docking task, for both 2D and 3D visualization conditions. This study showed that direct-touch shortens completion times, but indirect interaction improves efficiency and precision, and this is particularly true for 3D visualizations. The study also showed that users' trajectories are comparable for all conditions (2D/3D and direct/indirect). This tends to show that indirect RST control may be valuable for interactive visualization of 3D content. To illustrate this finding, we present a demo application that allows novice users to arrange 3D objects on a 2D virtual plane in an easy and efficient way.

Index Terms: H.5.2 [Information Interfaces and Presentation]: User Interfaces—Input devices and strategies

1 INTRODUCTION

Recent work on multi-touch surfaces has focused on direct interaction, where the input space and the visualization space are co-located. It is commonly argued that direct-touch favors fast and straightforward interaction. On the other hand, indirect multi-touch interaction has some substantial advantages that have been highlighted in the recent literature (eg. [7][9]). In particular, indirect multi-touch design is not guided by co-location concerns, which seems particularly interesting for interactive 3D visualization. The goal of this paper is to understand how directness influences user performance in 2D and 3D spatial contexts. In our approach, we do not attempt to use multi-touch gestures to control the 6 degrees-of-freedom (DOFs) that can be implied in 3D interaction. Instead, we focus on the well-known *RST* technique that allows the simultaneous control of *Rotations*, *Scaling*, and *Translations* from two finger inputs [3][6]. In particular, we wanted to investigate if users were able to interact on 3D-oriented planes as they do with the standard 2D RST technique operated on the screen-plane. To complement the study we developed a demo application, which is illustrated in Figure 1.

A lot of research was conducted in past years to investigate and evaluate multi-touch interaction techniques in 2D spatial context. For example, Latulipe et al. [4] showed that symmetric bimanual interaction outperforms the single mouse interaction during image alignment tasks, when users have to manipulate multiple degrees of freedom simultaneously. Moscovich and Hughes [7] analyzed two-touch interaction with one and two hands for position and orientation control. They showed that performance of two hands was better when separating interaction tasks. Schmidt et al. [9] compared direct and indirect multi-touch interaction on large surfaces.

*e-mail: sebastian.knoedel@googlemail.com

†e-mail:Martin.Hachet@inria.fr

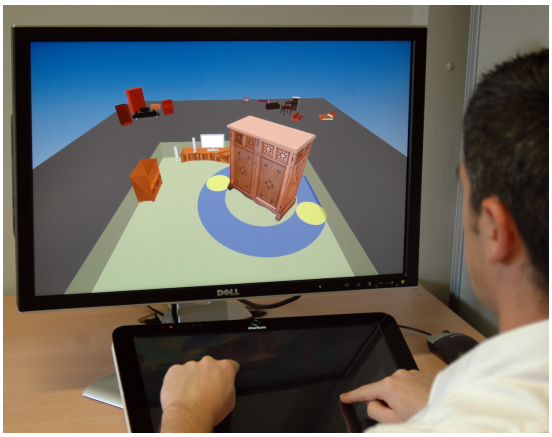


Figure 1: Indirect control of the RST technique in a 3D furnitures manipulation scenario.

They analyzed completion times of dragging and resizing operations of 2D content during a docking task experiment in direct and indirect conditions. In contrast to their approach, we consider rotation operations as well and conducted further measurements on efficiency and precision to draw conclusions on users coordination. In addition, we are interested in both 2D and 3D visualization setups.

First attempts to create applications that allow the control of 3D content through multiple touch input were made by Zeleznik et al. [10]. In recent years, researchers started to combine the multi-point input technology with interactive 3D graphics. Hancock et al. [2] proposed different approaches using direct and indirect multi-touch interaction techniques for tabletop surfaces to manipulate 2D and 3D content. Martinet et al. have studied the effect of DOF separation in a 3D manipulation task on a direct multi-touch display [5]. In their work, Reisman et al. [8] proposed co-location constraints for direct-touch object manipulation. The Cubtile has been designed by de la Rivière et al. [1], where users interact with 3D content from an indirect multi-touch box-shaped device.

In our approach, we focus on several aspects that were not addressed by previous work. In particular, we explore how well the RST technique adapts to the 3D visualization space, by looking closely at user performance over direct and indirect conditions. Our overall hypothesis was that indirect interaction may be valuable when visualizing 3D data.

2 USER STUDY

2.1 Design

We designed an experiment where the task consisted in aligning one source object with a target object. To complete the task, subjects had to control rotations (1 DOF), scaling (1 DOF), and translations (2 DOF), on either the screen plane (2D) or on a 2D plane oriented in 3D space (3D). The source object to manipulate was a blue rectangle in the 2D scenario and a blue box in the 3D scenario

(see Figure 2). In all cases both objects were textured with an arrow indicating one distinct orientation. The source object appeared always in the center of the scene at the same position. Initial size and orientation changed randomly. The grey target object was placed with random position, size and orientation around the source object. During the experiment, the viewing angle was set to 45° to ensure a good visualization of the 3D data. Four red spheres on the upside corners of the target highlight the tolerance areas. They represent the threshold to successfully complete the task. When aligning one corner of the source object correctly with the corresponding corner of the target object in range of the threshold, the color of the sphere turned from red to green. When correctly aligned, the source object had to remain 0.7 seconds to complete the task.

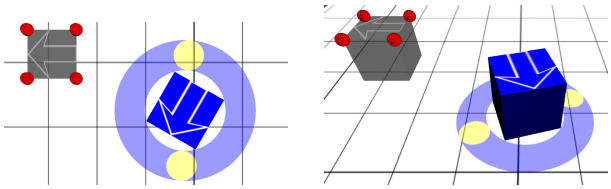


Figure 2: Docking task. The source object (blue) has to be aligned with target object (grey transparent), in 2D (a) and in 3D (b).

We used a within-subject repeated-measures experiment. The independent variables were the visual representation (2D/3D) and the separation condition (direct/indirect). Note that for the 3D indirect condition, the fingers' movements on the touch sensor were directly mapped to cursor movements on the 2D virtual plane that is oriented in 3D space. Participants performed for each of the four conditions 4 repeated tests, which consisted of 8 randomly shuffled trials. Each subject conducted 128 trials in total. The order of presentation of the four conditions was balanced according to a Latin square. Figure 3 summarizes the experimental conditions.

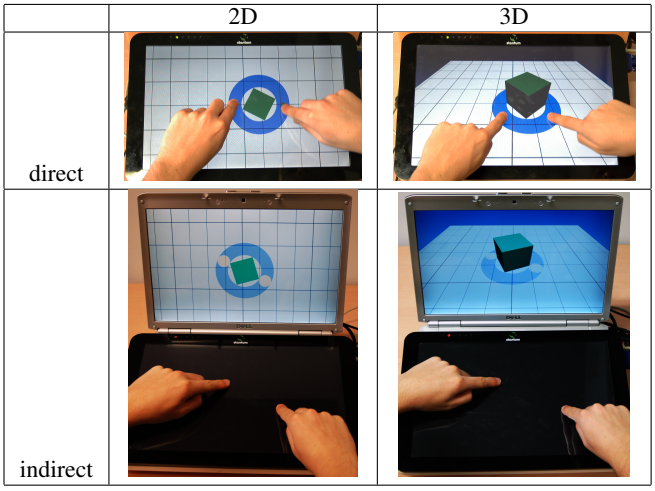


Figure 3: Experimental conditions.

24 subjects (8 women, 16 men) participated in the experiment. The subjects used index fingers of both hands to interact with the interface. They were regular computer users, but had no previous experience using a multi-touch interface. Each subject needed approximately 30 minutes to perform the task and to fill out a short questionnaire.

During our study we used a Stantum SMK resistive multi-touch screen as haptic sensor. The display has a size of 15" with a resolution of 1280x800 for direct interaction and a touch repositioning accuracy of less than four pixels or less than 0.5 mm. For indirect interaction we used a Dell Notebook display with equal size and identical resolution. All content was rendered with constant 60 frames per second synchronized with the display refresh rate under all four conditions.

We measured efficiency of the trajectories, accuracy, and completion time. For efficiency, we extended the *inefficiency* measure introduced by Zhai and Milgram [11]. Inefficiency is the ratio between the distance traveled by users d_u and the optimal distance d_o between current position and target; defined as $I = \frac{d_u - d_o}{d_o}$. Inefficiency can be computed for translation, rotation and scaling, as well as for 2 dimensional spaces (e.g. I_{TR} for translation-rotation) and 3 dimensional spaces (ie. I_{RST} for rotation-scaling-translation). The final I_{RST} measure is the ratio between the user's curve length l_{RST} and the length of the optimal curve l_o, $I_{RST} = \frac{l_{RST} - l_o}{l_o}$. To measure accuracy, we use a *target reentry* metric. Completion time is given by the duration starting from the time where the subjects touch the multi-touch surface, to the time when the task is successfully completed.

2.2 Quantitative Results

We performed two-way ANOVAs on I_{RST}, target reentry, and completion time, which showed the following effects. Figures 4 illustrates the obtained means.

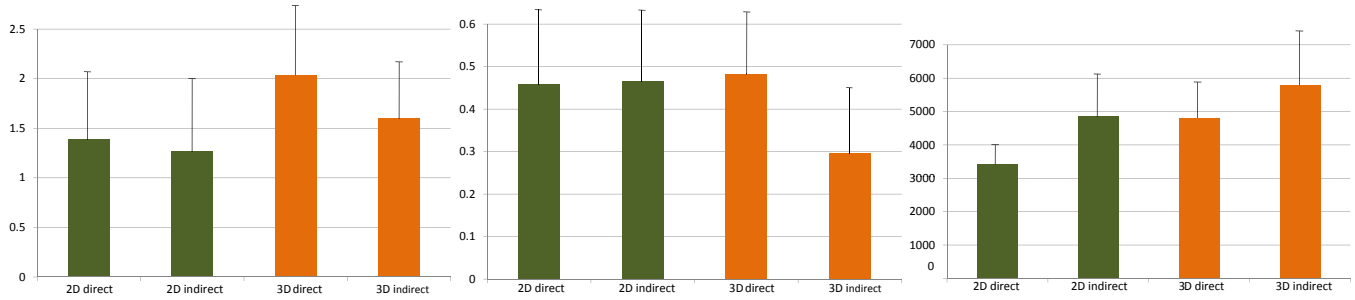
The subjects performed the task 27.1% more efficiently in 2D than in 3D (2D:1.322, 3D:1.814, $F_{1,23} = 34.243, p < .0001$). Similarly, they were 22% faster (2D:4125ms, 3D:5290ms, $F_{1,23} = 36.125, p < .0001$). On the other hand, no significant difference allows us to confirm an effect on precision.

If we compare direct and indirect modes, we see that indirect interaction is 14.4% more efficient (1.429) than direct interaction (1.708) ($F_{1,23} = 6.576, p < .05$), and subjects are 21.1% more precise (direct:0.47 indirect:0.38, $F_{1,23} = 5.860, p < .05$). This indicates that indirect interaction does not lead to less efficient/precise interaction. Instead users interact even more efficiently and more precisely in indirect mode. Moreover, significant interaction effects show that the influence of the separation mode is bigger in 3D than in 2D (I_{RST}: $F_{1,23} = 6.439, p < .05$, target reentry: $F_{1,23} = 13.940, p < .05$).

Regarding completion time, direct interaction (4115ms) is 22.4% faster than indirect interaction (5300ms) ($F_{1,23} = 37.858, p < .0001$). An interaction effect shows that the influence of distance on completion time is smaller in 3D than in 2D ($F_{1,23} = 4.345, p < .05$).

One important premise of this paper is to investigate user performance in 3D conditions. Therefore, we analyzed interaction effects on precision and efficiency between direct and indirect manipulation modes in 3D. A t-test on I_{RST} reveals a significant effect ($t = 3.607, p < .05$) showing that in 3D indirect users are 21.6% more efficient (1.595) than in 3D direct (2.034). A similar significant effect can be observed for target reentry (direct:0.481, indirect:0.296, $t = 4.735, p < .0001$). Regarding completion time, direct interaction remains significantly faster (4820ms) than indirect interaction (5760ms) ($t = -4.345, p < .05$). Furthermore, our questionnaire shows that 14 out of 24 subjects prefer to manipulate objects from a distance (Wilcoxon Z Test: $-2.521, p < .05$), when interacting in 3D.

To evaluate the learning effect in the 3D indirect condition, we conducted one-way ANOVAs on I_{RST} and target reentry. Both show no significant difference between the mean values of each consecutive tests during the study (respectively, $p > .500$, $p > .900$). Consequently, we cannot show any performance improvement over time. Regarding completion time, our data shows that the completion of the first 8-trials are more time consuming than the following 8-trials tests. No improvement operates starting from the second



(a) I_{RST} (b) Target Reentry (c) Completion Time

Figure 4: The total mean values for inefficiency, target reentry, and completion time for all four conditions.

test. A similar behavior can be observed for the other conditions (2D/3D, direct/indirect).

2.3 Discussion

The results presented above allow us to better understand the influence of direct/indirect interaction under 2D and 3D visualization on user performance. In particular, this study highlights that, although direct touch is faster, indirect interaction is more efficient and more precise for RST docking tasks. This can be explained by the fact that dissociating the visual space and the interaction space releases the co-location constraints. We showed that this effect was bigger in 3D than in 2D. Consequently, we believe that indirect interaction may be more suitable for interactive visualization of 3D data. Moreover, the results of the study showed that, with direct interaction, the speed gain is lower in 3D than in 2D.

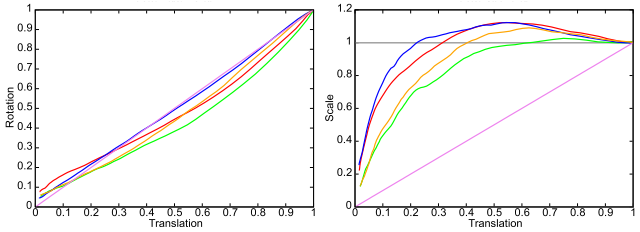
In our experiment, we set the orientation of the virtual ground to 45° as a compromise between a good 3D visualization and a convenient configuration for both direct and indirect 3D interaction modes. We can presume that, with orientations of the ground closer to the direction of the viewing direction, the differences of performance between the 3D direct mode and the 3D indirect mode would have been bigger. Indeed, with direct-touch, small finger movements on the screen can project on large distances in the 3D environment, resulting in a loss of precision.

With 3D indirect, the mapping between the multi-touch surface and the virtual ground seems to work well. The analysis of performance evolution over time did not show learning effect. This might indicate that learning of the technique is very fast and takes place during the first trials, or that improvements in performance with time are not large enough to be of statistical or practical significance. This has been confirmed by the subject's comments we obtained from the post-experiment questionnaire. The subjects rated their experiences on a 5 point Likert scale (1 fully disagree and 5 fully agree). They rated the indirect 3D multi-touch interaction technique as easy to use with an average score of 4.4. Moreover, the subjects stated that they were able to solve the task easily with an average score of 4.0. These scores are almost similar to the scores obtained for the 2D direct approach. 2D direct multi-touch interaction currently knows many successes, in particular because users are able to complete interaction tasks in a very easy and convenient way. The experiment we have conducted tends to show that a similar benefit can be obtained under 3D visualizations.

2.4 Coordination

Coordination represents the ability of users to control several DOFs at the same time. Since one key component of multi-touch interaction is the possibility to manage several DOFs simultaneously, it is interesting to understand how coordinated users' movements are. The study of the shapes of the obtained inefficiency curves teach us about coordination. We computed the mean curves in the translation-rotation space I_{TR}, the translation-scaling space I_{TS}, the rotation-scaling space I_{RS}, as well as in the rotation-scaling-translation space I_{RST}. Figure 2.4 (a) shows that the I_{TR} curves, in all conditions, follow the characteristic of the optimal curve (purple). Consequently, translations and rotations seem to be controlled in a coordinated way, independently of the visualization mode and the separation condition.

In contrast, the characteristic of I_{TS} and I_{RS} mean curves (Figure 2.4 (b)) show that users tend to separate scaling from the other DOFs. In both cases, users perform rotation and translation first, and then scale the object. The same behavior is observed in any conditions. This tends to show that users interact with 3D data at-a-distance in a similar way they would do with 2D data displayed on a multi-touch screen.



(a) Translation-Rotation, I_{TR} (b) Translation-Scaling, I_{TS}

Figure 5: Mismatch curves for all four conditions and optimal curve (purple).

3 Demo Application

We designed a demo application benefiting from RST multi-touch interaction where users manipulate 3D objects on a 2D plane from a distance. This application, dedicated to the general public, can be seen as an extension in 3D of the standard 2D multi-touch image viewer. Such an application can be valuable in many domains such as entertainment and video games, cultural heritage, education, automotive industries, and so on. In this work, we used a scenario where users can move, turn and resize furniture to arrange their living room by sliding their fingers on the multi-touch surface (see Figure 1).

Selection. Since visualization and interaction are decoupled, a remote selection mechanism must be introduced. In our approach, users are able to move a standard cursor on the screen -as known from desktop applications- by touching the tactile sensor with a unique finger. When touching the tactile surface with their second

finger, users select the pointed object. The application moves from a *cursor mode* to a *manipulation mode*. A blue circle appears below the selected object, and the cursor corresponding to the user's fingers are highlighted by way of yellow disks. The application stays in manipulation mode until a single finger movement is detected. Hence, users are able to release their fingers to manipulate comfortably the selected object. To select a new object, they only need to slide one finger on the multi-touch surface to make the cursor appear.

Global camera control. In addition to selection and manipulation, RST multi-touch interaction can be used for global camera control, similar to [10]. After having selected the ground surface, users are able to zoom in and out, drag the world and orbit around it in a similar way they do when manipulating a single object. The previously described RST technique is used to transform the global scene, which results in the modification of the camera viewpoint. Finally, a virtual slider, located on the left side of the multi-touch surface, allows changing the orientation of the ground plane (pitch angle).

First user feedback. We asked colleagues to test this demo application to obtain a initial feedback. The general usability appeared to be good. Users had a very short experience with multi-touch surfaces, and they had never experimented our application before. They found the application very easy to use. They were able to position the 3D objects where they wanted, and they felt comfortable using the selection mechanism. They easily switched between the object manipulation mode and the camera control mode. They reported that they particularly appreciated the possibility to change the view to the world in a fast and easy way. After a short time of use, they were able to interact efficiently with the application. This confirmed the results we had from the user study described above. Of course, such an interface does not allow the control of all the DOFs that could be manipulated in the 3D space (eg. roll and vertical translations). It rather provides a simple interface for manipulation and observation of 3D objects resting on a virtual ground. Such a 3D interaction paradigm may be very efficient for many 3D applications where gravity plays an important role.

A direct touch approach could be used for the same application. It may appear as more "natural", and the selection process can be sped up. On the other hand, our study and our experience with the demo application showed that users performance and general usability of the 3D-extended RST technique is good when used in indirect mode. Consequently, similar applications can benefit from the advantages of indirect-touch interaction (e.g. for stereoscopic and large display visualization) and can still be used easily and efficiently by inexperienced users.

4 Conclusion and Directions to Future Work

2D applications based on multi-touch input are now well-known and accepted by the general public. Such applications are changing the way we interact with digital content. On the other hand, interaction with 3D content on multi-touch surfaces is not as mature. In particular, the difference between the 2D structure of the input surface and the 3D structure of the visualized data leads to a more complex interaction process. In this paper, we have shown that an adapted version of the RST technique allows users to benefit from multi-touch input for constrained manipulations of 3D objects. By moving cursors on a virtual ground plane, users can move, rotate, or scale 3D objects in a very convenient way.

The user study showed that indirect interaction favors efficiency and precision for a RST docking task. This is particularly true for 3D visualization, because users can separate their finger movements from the 3D visualization. Consequently, we argue that users can benefit from indirect interaction when dealing with 3D data. We illustrated such a configuration with a demo application, which shows the potential of such an interface.

In this paper, we focused in interaction movements on a virtual ground. It would be interesting to see if the same approach could be adapted to other surfaces. For example, we would like to know if the semantic link between the fingers movements on the sensor and the resulting actions on the target surface still works if the latter has an arbitrary orientation (eg. a wall of a room that does not face the camera). One could also imagine different kind of mappings with curved surfaces. These investigations will be part of our future work.

The 3D-oriented RST technique has a limited range of actions. However, we are convinced that such an interface is very valuable compared to the existing mono-point approaches. It allows novice users to well perceive 3D objects by manipulating them easily. In the future, we want to explore new directions for 3D multi-touch interaction. In particular, it would be interesting to investigate good mappings between the multi-touch input and other DOFs in the 3D environment. For example, one can imagine to control non-homogeneous scaling, or non-canonical rotations. Many other 3D tasks are still to be studied. Multi-touch brings new challenges to research in the scope of 3D user interfaces. It also opens many possibilities for future 3D applications.

5 Acknowledgments

This work was supported by the ANR project ANR-09-CORD- 013 InSTInCT - http://anr-instinct.cap-sciences.net.

References

[1] J.-B. de la Rivière, C. Kervégant, E. Orvain, and N. Dittlo. Cubtile: a multi-touch cubic interface. In *VRST '08: Proceedings of the 2008 ACM symposium on Virtual reality software and technology*, pages 69–72, 2008.

[2] M. Hancock, O. Hilliges, C. Christopher, D. Baur, and S. Carpendale. Exploring tangible and direct touch interfaces for manipulating 2d and 3d information on a digital table. In *Proceedings Interactive Tabletops and Surfaces 2009*, pages 85–92, 2009.

[3] G. Kurtenbach, G. Fitzmaurice, T. Baudel, and B. Buxton. The design of a gui paradigm based on tablets, two-hands, and transparency. In *CHI '97: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 35–42, New York, NY, USA, 1997.

[4] C. Latulipe, C. S. Kaplan, and C. L. A. Clarke. Bimanual and unimanual image alignment: an evaluation of mouse-based techniques. In *UIST '05: Proceedings of the 18th annual ACM symposium on User interface software and technology*, pages 123–131, 2005.

[5] A. Martinet, G. Casiez, and L. Grisoni. The effect of dof separation in 3d manipulation tasks with multi-touch displays. In *Proceedings of the 17th ACM Symposium on Virtual Reality Software and Technology*, VRST '10, pages 111–118, 2010.

[6] T. Moscovich. Multi-touch interaction. *Conference on Human Factors in Computing Systems*, pages 1–4, 2006.

[7] T. Moscovich and J. F. Hughes. Indirect mappings of multi-touch input using one and two hands. In *CHI '08: Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, pages 1275–1284, 2008.

[8] J. L. Reisman, P. L. Davidson, and J. Y. Han. A screen-space formulation for 2d and 3d direct manipulation. In *UIST '09: Proceedings of the 22nd annual ACM symposium on User interface software and technology*, pages 69–78, New York, NY, USA, 2009. ACM.

[9] D. Schmidt, F. Block, and H. Gellersen. A comparison of direct and indirect multi-touch input for large surfaces. In *INTERACT '09: Proceedings of the 12th IFIP TC 13 International Conference on Human-Computer Interaction*, 2009.

[10] R. C. Zeleznik, A. S. Forsberg, and P. S. Strauss. Two pointer input for 3d interaction. In *SI3D '97: Proceedings of the 1997 symposium on Interactive 3D graphics*, pages 115–., 1997.

[11] S. Zhai and P. Milgram. Quantifying coordination in multiple dof movement and its application to evaluating 6 dof input devices. In *CHI '98: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 320–327, 1998.