Shallow-Depth 3D Interaction: Design and Evaluation of One-, Two- and Three-Touch Techniques

Mark Hancock, Sheelagh Carpendale
University of Calgary
Calgary, AB, Canada
{msh,sheelagh}@cs.ucalgary.ca

Andy Cockburn
University of Canterbury
Christchurch, NZ
andrew.cockburn@canterbury.ac.nz

ABSTRACT
On traditional tables, people frequently use the third dimension to pile, sort and store objects. However, while effective and informative for organization, this use of the third dimension does not usually extend far above the table. To enrich interaction with digital tables, we present the concept of *shallow-depth 3D* – 3D interaction with limited depth. Within this shallow-depth 3D environment several common interaction methods need to be reconsidered. Starting from any of one, two and three touch points, we present interaction techniques that provide control of all types of 3D rotation coupled with translation (6DOF) on a direct-touch tabletop display. The different techniques exemplify a wide range of interaction possibilities: from the one-touch technique, which is designed to be simple and natural, but inherits a degree of imprecision from its simplicity; through to three-touch interaction, which allows precise bimanual simultaneous control of multiple degrees of freedom, but at the cost of simplicity. To understand how these techniques support interaction in shallow-depth 3D, we present a user study that examines the efficiency of, and preferences for, the techniques developed. Results show that users are fastest and most accurate when using the three-touch technique and that their preferences were also strongly in favour of the expressive power available from three-touch.

Author Keywords
Shallow-depth 3D, tabletop display, direct-touch, rotation and translation.

ACM Classification Keywords
H5.2 [Information interfaces and presentation]: User Interfaces - Interaction styles, Input devices and strategies.

INTRODUCTION
In this paper, we consider the concept of *shallow-depth 3D* – full 3D visuals with full 3D interaction, but extremely limited depth – as a potential interaction space. This work is motivated by the new generation of hardware and software that more closely emulates physical workspaces. Tabletop

displays, for example, have the potential to bring the advantages of electronic media to the type of interactions that occur over traditional desktops, and software environments like 'BumpTop' [1] greatly enhance the reality of interaction through physics modelling. But these environments, and others like them, are limited in that they typically allow only a single point of cursor or pen-based contact for manipulating objects from a single perspective through a single interaction plane. Compared to the rich means available for controlling physical objects, this single point can be limiting.

Several hardware technologies allow multiple concurrent points of contact [2, 5] and researchers have used them to create direct manipulation techniques that naturally emulate 2D rotation, translation, and scaling [9, 6, 13]. 3D manipulations, however, have not yet been fully explored.

We focus on shallow-depth interaction in the z-plane for two reasons: first, interactions on physical desktops take place within a shallow-depth field (e.g. riffling, sorting and manipulating piles, and rotating or flipping objects on the surface); second, current desktop graphical user interfaces are similarly limited to a shallow-depth field. We argue in this paper that providing users with shallow-depth 3D capabilities allows for a more engaging and rich experience.

To empirically explore shallow-depth 3D as an interaction space, we consider the task of moving and rotating a small 3D object (e.g. a cube) across a tabletop (see Figure 1). We first present design guidelines for direct-touch 3D interaction. Next we discuss alternative candidate interaction techniques for supporting these manipulations using one, two and three points of contact, formally demonstrating how two-dimensional surface interactions can be used to directly manipulate shallow-depth 3D objects. We then describe a usability study that compares the speed and accuracy of the techniques as well as the users' subjective perceptions of them. In closing we discuss the implications and suggest two alternative techniques based on the results of this study.

RELATED WORK
Almost all windowing interfaces currently use shallow 3D effects to support interaction. The layering and shadowing effects both enhance the visual appeal of the interfaces and provide a natural metaphor for switching documents and workspaces into and out of focus. Some commercial interfaces further extend the 3D effects, using animations to clarify feedback effects such as distorting windows and icons
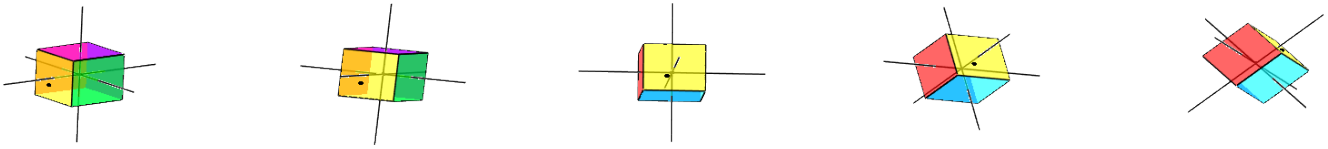
Figure 1. A sequence of motion using one-touch interaction in shallow-depth 3D. The black dot represents the point of contact of the user's finger.

to show the relationship between pre- and post-action states (e.g. Mac OS®X). Researchers are also investigating problems and solutions that arise from moving between layers on the desktop. Dragicevic [3], for example, describes 3D visuals of dog-ears, folding and shuffling to make working with overlapping windows more intuitive. Agarawala and Balakrishnan's [1] 'BumpTop' wholeheartedly adopts the emulation of reality on the desktop, using both rich 3D visuals and physics modelling to enrich interaction – objects can be piled on top of one another or flipped onto their backs; objects can be thrown at others, and the visual effects of collisions depends on their mass and velocity. The reality of the lustrous environment, however, is hindered by its constraint to a single point of interaction through a stylus input device. For comparison with the reality it attempts to emulate, though, consider the awkwardness of manipulating objects on your physical desk using only one index finger.

For collaborative tabletop displays, a variety of 3D effects have been proposed. Ståhl and Lundberg's [14] tabletop 3D virtual pond floats items in use to the surface and allows items to sink when they are no longer in active use. The Lumisight table [10] and Nakashima et al.'s 3D table [11] provide up to four users with a coherent view of a 3D image in the centre of the display. While these systems are capable of rich 3D visuals in a collaborative setting, they do not fully address the interaction with these 3D models. Furthermore, these systems require a very large tabletop to achieve a small central 3D display.

In the study of high degree-of-freedom (DOF) input devices, there has been a general consensus about the separability of rotation and translation. It is widely believed that input is superior if these are kept separate. Frölich et al. [4] provide a good discussion of this phenomenon. They both argue and demonstrate empirically that the separation into "forceless isotonic rotational input" and "force-requiring elastic translational input" is key to a good design of a 6DOF input device. However, other researchers suggest that rotation and translation are *not* separable in the human mind [7, 15]. Studies of 2D interaction techniques for 2D tabletop interfaces, such as RNT [9], which require only 3DOF, tend to confirm that users typically ignore this difference and that integration of rotation and translation is essential for the support of communication in a collaborative setting [8, 9, 6]. Although these two claims appear contradictory, we argue that rotation and translation can be separated, but performed simultaneously and that this provides a natural interface for communication without sacrificing performance.

3D TABLETOP INTERACTION DESIGN GUIDELINES

Our eventual goal is to enable the existing freedom of object manipulation available on physical tables within digital desktops. In this endeavour, we attempt to combine the benefits of both 3D interaction and digital tabletops. In order to be successful, we need to take care in the design of new techniques to support interaction. We suggest the following design guidelines for interaction on tabletop displays in 3D:

- *Provide More Degrees of Freedom*: Interactions including flipping of objects, storage and communication through small adjustments of objects become possible by allowing full rotation and translation in three dimensions.

  - *Simultaneity of Rotation and Translation*: In the real world, people are capable of simultaneously activating a combination of muscles to perform a single action that both moves and rotates an object. Similarly, a 3D tabletop interface should allow users to simultaneously rotate and translate an object.

  - *Independence of Rotation and Translation*: Along the same lines, users should be able to activate rotation and translation using distinct actions, in the same way that different muscle groups are used to perform these actions in reality. Thus, users could combine these actions at the cognitive level instead of combining them in a potentially awkward way through the interaction technique itself.

- *Provide Connected Manipulation*: Connected manipulation is direct manipulation in which the user maintains a constant visual and physical connection with the object *throughout the entire interaction*. We distinguish this from direct input – when input space and display space are superimposed – by specifically requiring that the object being manipulated, not just the display space, remain in physical contact with the input mechanism. This guideline is also important when manipulating 2D objects on a digital table [6, 8], but we emphasize its importance in 3D here, as it may be more tempting to ignore this constraint when more DOF are required.

- *Prevent Cognitive Disconnect*: By allowing users to interact directly with 3D objects, one should avoid actions that (though possible with technology) do not conform to what users expect. For example, since it is impossible to push one's finger through the display, limiting the depth of interaction maintains that expectation and prevents disconnect. Furthermore, traditional tabletops offer interaction on the surface of the table and the space between the top of the table and the user; however, most interaction takes place in the first few inches. Limiting interaction to a small finite z depth places a virtual surface just below the actual display, providing a similar few inches for 3D interaction.

- *Provide Appropriate 3D Visual Feedback*: 3D visual cues can make interactions more familiar, since they relate closely to the physical world.
  - *Provide Appropriate Shading*: Tabletops afford interaction from all sides of the table. This can pose problems in recognizing concavity and convexity because in the absence of other cues, lighting will be assumed to originate from over one's shoulder [16], causing a button to appear depressed from one side of the table but not the other. Use of full 3D projections with additional depth cues such as shape and cast shadows can reduce this effect.
  - *Consider Parallax*: As the z depth increases a single perspective projection can appear different to users at different locations. Even for a single user, parallax can occur from left to right, simply due to the size of the display and the proximity of the user.
- *Support Many Identifiable Points of Contact for **each** user at the table*: Current tabletop display hardware provides a mosaic of supporting technology. Some technologies allow for a large number of points of contact, without identifying information [5], and others provide identifiable input for a single point of contact for a small number of users [2]. In order to fully support direct-touch 3D interaction for multiple users on tables, the hardware needs to support identification of not only *where* a finger is touching, but also *which finger* of *which user* is touching.

DIRECT-TOUCH 3D INTERACTION TECHNIQUES

We have designed three new direct-touch interaction techniques for manipulating 3D objects on a tabletop display. These designs were in part informed by our suggested guidelines, but mostly have helped to generate them.

On the two-dimensional surface of the table, each point of contact provides two degrees of freedom of *input*. It is possible to create interaction techniques that vary from being essentially 2D input (one touch point and 2DOF), through being 2D+ input (two touch points and 4DOF), to being fully 3D input (three touch points and 6DOF). The shallow-depth 3D *output* we wish to provide has the following five degrees of freedom:

- x & y – the position on the surface of the table
- yaw – object rotation about the z-axis (planar)
- roll – object rotation about the y-axis (side-to-side)
- pitch – object rotation about the x-axis (front-to-back)

Note that we also describe how z (movement into the table) could be manipulated in some cases, but this has (purposefully) not been implemented. Future designs could apply this z movement to other degrees of freedom (e.g. scaling).

One-Touch Input
We can achieve 5DOF movement with a single point (2DOF input) by extending the RNT algorithm [9] into the third dimension. Instead of an explicit rotation about the axes (yaw, pitch and roll), the axis of rotation can be determined from the point of contact (see Figure 2).
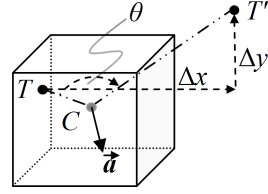

Figure 2. One-touch interaction in shallow-depth 3D.

Algorithm

Let $T = (x, y)$ and $T' = (x', y')$ be the initial and final points of contact.
Let $\overline{T} = (x, y, z)$, where z is the value in the z-buffer at (x, y) (i.e. the z-value of the point on the object "nearest" to the point T at the surface of the table).
Let $\overline{T'} = (x', y', z)$ (using the same z-value as $\overline{T}$).
Let $C = (x_c, y_c, z_c)$ be the initial centre of the object in the 3D scene.

$$\Delta x = x' - x, \Delta y = y' - y$$
$$\vec{a} = \overrightarrow{CT} \times \overrightarrow{CT'}$$
$$\theta = \angle \overline{T} C \overline{T'}$$

where, $\vec{a}$ is the axis of rotation and θ is the angle of rotation about that axis.

Description

While the discrepancy in mapping between input degrees of freedom (2DOF) and output degrees of freedom (5DOF) is high, the actual action feels quite natural (see Figure 1). Touching a point on the cube works like a sticky finger in that the contact point will rise toward the surface and the leading direction, causing the cube to rotate in x, y, and z until the contact point is as close to the surface and the lead direction as the shape of the cube will allow. Rotating the chosen side to the surface merely involves touching that side and dragging. This can require a re-touch for an initially occluded side.

Despite the fact that this technique provides the ability to rotate and translate a 3D object to any position and orientation, it is common for users to want to perform more constrained interaction, such as translation alone or planar rotation. We provide this ability through dedicated areas on the object. For polygonal objects, a circular area about the centre of each face is reserved for translations and a doughnut-shaped region around that circle is reserved for planar rotations (using the 2D RNT algorithm). For non-polygonal objects, a more abstract central location can be chosen on some surface of the object, about which the circle and doughnut shapes can be drawn.

Two-Touch Input
Five or six degrees of freedom of output can be achieved using only two points of contact (4DOF input). The first point of contact can use the RNT algorithm [9] to achieve both translation in x and y as well as yaw. The second point can be used to specify pitch and roll (see Figure 3). If z motion is desirable, this can be manipulated according to the change in distance between the two points.
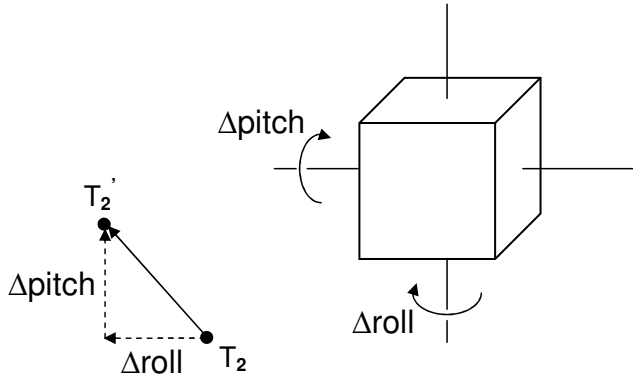
Figure 3. Using the two-touch interaction, users can perform 2D rotations and translations with the index finger of the dominant hand using the RNT algorithm while simultaneously performing pitch and roll rotations (shown above) with a second finger on the non-dominant hand.
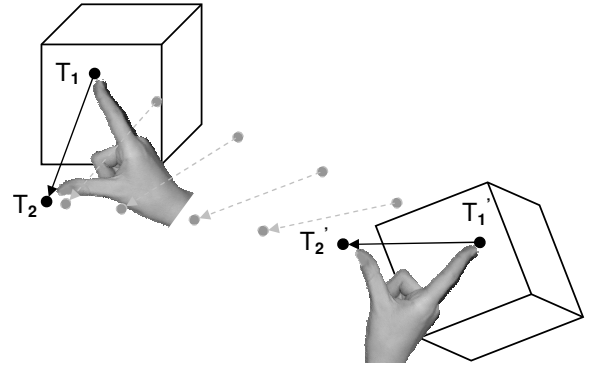

Figure 4. Using three-touch interaction, users can perform a simultaneous translation and rotation on the surface of the table, as shown here. The user can also simultaneously rotate the object in pitch and roll with a finger on the non-dominant hand.

Algorithm
Let $T_i = (x_i, y_i)$ and $T'_i = (x'_i, y'_i)$ be the initial and final points for the ith point of contact, where $i \in \{1, 2\}$.
Let $C = (x_c, y_c, z_c)$ be the initial centre of the object.
Let $C_{xy} = (x_c, y_c)$ (the projection of C in the 2D surface).

$$\Delta x = x'_1 - x_1, \Delta y = y'_1 - y_1$$

$$\Delta z = |\overrightarrow{T'_1 T'_2}| - |\overrightarrow{T_1 T_2}| \text{ (if desired)}$$

$$\Delta yaw = \angle T_1 C_{xy} T'_1 \text{ (about } T_1)$$

$$\Delta roll = K_1(x'_2 - x_2), \exists K_1 \in \mathbb{R}$$

$$\Delta pitch = K_2(y'_2 - y_2), \exists K_2 \in \mathbb{R}$$

In our user study, T_1 was provided through the index finger of the dominant hand and T_2 through the index finger of the non-dominant hand. However, the technique is not limited to this configuration; other sensible configurations include reversing these two fingers or using the index finger and thumb on the same hand.

Description
This technique provides easy causal movement coupled with rotation that maintains the vertical orientation of the object's projection. If the vertical orientation needs adjusting, for example if the right side of a cube is not at the surface, it can be adjusted with a finger on the non-dominant hand.

As with the one-touch technique, it is often desirable to perform constrained translation-only movement. This is again provided at the centre of each face of a polygon or an abstract central location on the surface of a non-polygonal object.

Three-Touch Input
Our three-touch interaction technique maps 6DOF input to 5 or 6DOF output (see Figure 4). In this mapping, the first point of contact is used for translation, the second point for yaw about the first point, and the third point for pitch and roll about the centre of the object. The depth can be specified by the difference in distance between the first and second touch points. The order of the points can be specified either by taking the points in order of contact with the table or in a predefined order (if the source of each point is identifiable).

Algorithm
Let $T_i = (x_i, y_i)$ and $T'_i = (x'_i, y'_i)$ be the initial and final points for the ith point of contact, where $i \in \{1, 2, 3\}$.

$$\Delta x = x'_1 - x_1, \Delta y = y'_1 - y_1$$

$$\Delta z = |\overrightarrow{T'_1 T'_2}| - |\overrightarrow{T_1 T_2}| \text{ (if desired)}$$

$$\Delta yaw = \angle T_2 T_1 T'_2 \text{ (about } T_1)$$

$$\Delta roll = K_1(x'_3 - x_3), \exists K_1 \in \mathbb{R}$$

$$\Delta pitch = K_2(y'_3 - y_3), \exists K_2 \in \mathbb{R}$$

In our user study, T_1 and T_2 were provided through the index finger and thumb of the dominant hand, respectively. T_3 was provided through the index finger of the non-dominant hand.

Description
In principle this interaction is quite simple. For example, a single touch with one's index finger supports translation only, including one's thumb adds rotation around the z-axis and the addition of a finger from one's other hand provides the other two rotations.

In theory, three-touch allows the most efficient means for control because users can concurrently and independently manipulate all six degrees of freedom. However, there is a risk that this freedom may be confusing for users. Furthermore, both the two- and three-touch techniques may disconnect the object from the initial touch location upon rotation in pitch or roll. This disconnect may further confuse the user, creating an advantage to the one-touch technique. Hence, empirical comparison of the techniques is necessary.

USER STUDY
To better understand how people interact with these three (one, two and three finger) connected manipulation rotation and translation techniques, we conducted a study that compares them in terms of their speed, accuracy and the subjective preferences of the participants. Since these techniques vary considerably in interaction styles, conducting an empirical comparison can shed light on which balance of design tradeoffs are the most effective and satisfying for people.

For example, one-touch interaction is likely to be slow, but users may appreciate its simplicity and reliability; three-touch interaction may be fast if the participants can adapt to its comparative power and complexity, but they may report a higher cognitive load if it fails to be perceived as 'natural'.

Method

Participants
Twelve students (6 male, 6 female) from a local university participated in the study. Both national and international students were selected from a variety of disciplines. Five participants reported no prior experience with 3D gaming and seven reported some. The experience of these seven varied from once a year to four times a month. Ages ranged from 21 to 33 ($M = 26.3$, $SD = 3.8$). All participants were right-handed and no participant reported any colourblindness.

Apparatus
The experiment was performed on a front-projected 1024×768 pixel tabletop display using DiamondTouch [2] input with an 87 cm $\times$ 66 cm display area (12 pixels / cm). Multi-finger input was provided by attaching distinct Diamond-Touch sensors to both the index finger and thumb of a right-handed insulating glove. The third touch-point was provided with a regular DiamondTouch pad through the left hand. The display surface was 72 cm above the floor and participants were provided with an adjustable chair. An orthogonal 3D projection was used to render objects on the display. Objects were all full 3D objects but to provide the shallow-depth environment there was no movement in z. That is, objects could roll, tumble, and flip but the object's centre remained at a fixed z depth. Thus interaction was limited in all conditions to 5DOF. Software automatically logged the users actions and task times.

Procedure and Design
For each technique (one-, two-, and three-touch), participants performed three tasks in the same order. The order of techniques was counterbalanced between participants using a Latin square. Afterwards, each participant was asked to complete a questionnaire to provide both background and feedback about their experience. Participants were then interviewed by the experimenter.

The primary dependent measure in the two formally analysed tasks (tasks 1 and 2) was the task completion time. We additionally analysed data characterising how the users interacted with the techniques, including the time spent touching, translating and rotating the objects, and also the locations on the objects that the users touched.

Task 1: Passing
In order to determine a person's ability to use each technique for communication with other people, our first task required participants to pass a cube to one of three "virtual" people with a specific side of the cube facing upward and toward the virtual person (see Figure 5). This task was modelled after the study done on the 2D RNT rotation technique [9].
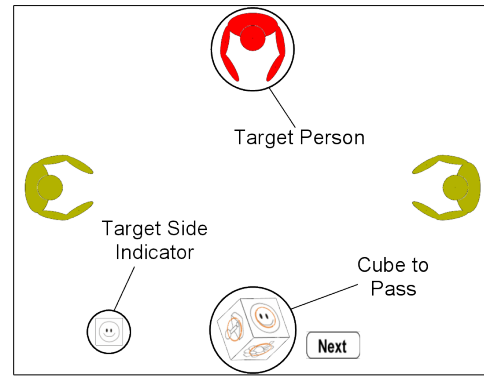
Figure 5. In the passing task, participants were asked to pass a cube to a target person with the target side facing up and toward the "virtual" user. The start position of the cube was close to the centre of the table.

Each side of the cube had a distinct grayscale icon. At the start of each trial, the cube was in the same location immediately in front of the participant, with the "top" side (a happy face) uppermost. Virtual participants were located to participants left, right and opposite. To start each trial an icon appeared on the screen and one virtual participant was indicated in red. The participant then matched that trial icon with one on the cube and passed the cube to the indicated virtual participant with the correct icon facing upward. The task was repeated six times – once for each side of the cube – for each target destination. Participants performed six practice trials each time they started with a new interaction technique.

Data from this task were analysed using a within-subjects analysis of variance for the following three factors:

- *Technique*: one-touch, two-touch, three-touch;
- *Destination*: left (40 cm), opposite (38 cm), right (40 cm);
- *Target-side*: top, bottom, left, right, back, front;

Task 2: Docking
To explore performance differences in the three techniques, we asked participants to complete a docking task. This task was a variation of the task developed by Zhai and Milgram [17] and used more recently to compare GlobeFish and GlobeMouse to other 6DOF techniques [4].

In this task, participants were asked to dock a tetrahedron inside another of equal size (see Figure 6). Spines around the
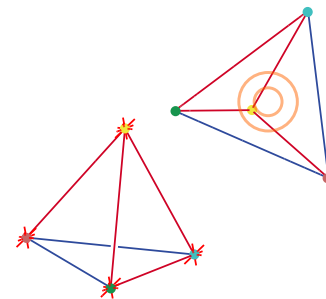
Figure 6. In the docking task, users were asked to dock a tetrahedron object (right) in another tetrahedron (left).
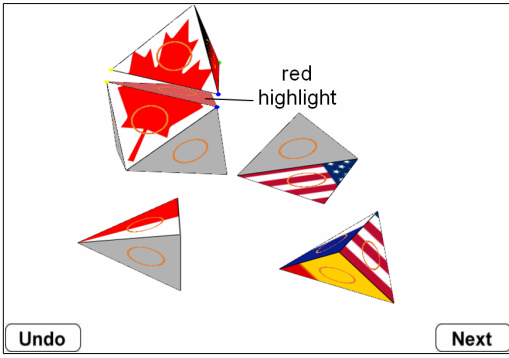
Figure 7. The puzzle task.

vertices were used to indicate docking tolerance. The vertices and edges of the tetrahedra were coloured to aide the participants in determining object orientation and the edges were haloed to aide in depth perception. When a given vertex was moved within target range, the vertex would change colour. Once all four vertices were in place for 700 ms, the source tetrahedron would disappear and the participant could begin the next trial by pressing the start button. Each trial had a 40 second time limit, after which the trial was abandoned and the next trial automatically began.

Trials were repeated for three levels of difficulty and for two levels of orientation. The levels of difficulty varied the size of tolerance bars at each vertex on the destination tetrahedron – easy trials had a 54 pixel tolerance, medium trials 38 pixels, and hard trials 23 pixels. The two levels of orientation allowed us to compare the techniques' support for planar rotations with more complex rotations – planar rotations used a 135° rotation about the z-axis, and complex rotations used and a 135° rotation about the x-y-z-axis.

Participants performed five repetitions of each combination of difficulty and starting orientation. Each time they began again with a new technique, participants performed six practice trials (each combination of difficulty and starting position was performed once).

Data from the docking task were analysed using a three-factor within-subjects analysis of variance on the factors:

- *Technique*: one-touch, two-touch, three-touch;
- *Difficulty*: easy, medium, hard.
- *Rotation*: planar, spatial.

Task 3: Puzzle
This task was used to examine how the participants' chose to use each of the techniques when completing a more realistic and less constrained task. Participants were asked to assemble a tetrahedron-shaped puzzle composed of four smaller tetrahedron shapes and a centre piece (see Figure 7). Participants performed this task once for each interaction technique. Although software logged the users' actions, data from this task was not formally analysed; our interest here was in observations of use, subjective preferences and comments about the techniques.

In summary, each participant was asked to do the following:

for each technique (in counterbalanced order):
- complete 18 passing trials in random order (6 sides × 3 destinations)
- complete 30 docking trials in random order (5 repetitions × 3 difficulties × 2 rotations)
- complete a puzzle task

The random ordering of passing and docking trials was different for each participant and technique.

RESULTS
Task Completion Times
Data from the task completion times violated Mauchly's test of sphericity for the repeated measures analysis of variance. We therefore report results using the Greenhouse-Geisser correction (influencing df, F and p values).

Task completion times (TCT) in both the passing and docking tasks showed the same trend, with users successfully exploiting the more expressive capabilities of the two- and three-touch interaction techniques. These results are summarised in Figure 8.

Passing Task
There was a marginally significant main effect of *technique* ($F_{1,16} = 3.4$, $p = .07$), with mean times reducing from 18.9 s ($SD = 2.8$) with one-touch, through 15.7 s ($SD = 2.2$) with two-touch, to 13.3 s ($SD = 2.0$) with three-touch; a 30% reduction in task time across the three conditions. Post-hoc pairwise comparisons only showed a significant difference between one-touch and three-touch techniques ($p < .01$). Despite the comparative efficiency of the three-touch technique, it is worth noting that even its mean task times were high—few tasks involving passing real objects would take this long, regardless of the level of precision required. We return to this issue in the discussion.

There was no significant effect of *destination* ($F_{1,18} = 0.06$, $p = .91$), nor were there significant interactions between it and the other two factors, suggesting that performance with the techniques is not substantially influenced by the direction of information sharing.
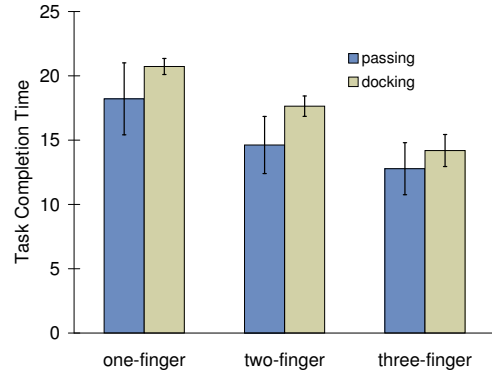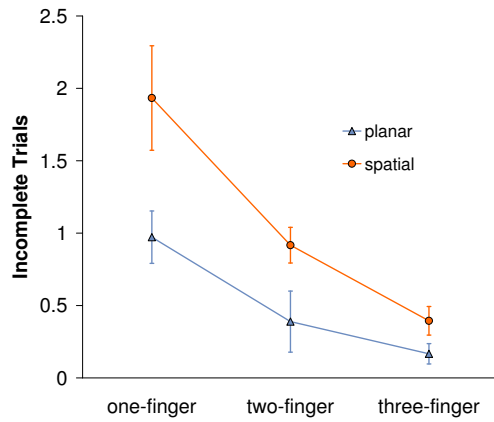

Figure 8. Mean TCTs for both the passing task and the docking task. In both cases, three-touch interaction is fastest, followed by two-touch, and one-touch is slowest.

Figure 9. There was a significant interaction between technique and required rotation for the number of incomplete trials. For one-touch interaction, the difference between planar-only trials and trials requiring spatial rotation was larger than for two- and three-touch interaction.

The *target side*, however, did have a significant effect on task performance ($F_{3,36} = 11.8$, $p < .001$). The mean time to attain the top-side target (10.6 s) was markedly lower than all others (bottom: 17.1 s; left: 17.2 s; right: 18.3 s; back: 17.3 s; and front: 15.5 s). This effect is explained by the lack of rotation necessary with the top side as the target. Such tasks, therefore, predominantly involved translation and planar rotation rather than the more complex spatial rotations required with the other sides. Post-hoc analysis showed pairwise differences ($p < .05$) between the top side and all other sides, and between the right and front sides. This latter difference is likely due to the combination of the facts that forward motion can more easily combine the required translation and rotation (advantaging discovery of the front side) and that our participants were right-handed, causing occlusion and disadvantaging trials involving the right side.

Docking Task
The results for the docking task showed similar trends to those for the passing task, but with stronger significance. Mean performance of the docking task with the three techniques improved significantly ($F_{2,19} = 14.2$, $p < .001$) as the number of touches increased from one to three. Means for the one-, two- and three-touch techniques were 20.1 s, 17.0 s and 14.3 s respectively (see Figure 8), showing a similar overall percentage improvement between one- and three-touch (29%) to that observed in the passing task. Post-hoc pairwise comparison showed significant differences between one-touch and both others ($p < .01$), and a marginal difference between two- and three-touch ($p = .06$).

As anticipated, there was a significant effect of *difficulty* ($F_{2,20} = 39.6$, $p < .001$), with means rising from 15.0 s on easy tasks, through 16.9 s on medium ones, and 19.6 s on hard tasks (post-hoc pairwise significant for all pairs at $p < .01$). Somewhat surprisingly, though, there was no technique × difficulty interaction ($F_{3,29} = 0.5$, $p = .77$). We had anticipated that one-touch may suffer more than the other techniques on high precision tasks because it does not allow independent manipulation of each degree of freedom, but the data did not support this hypothesis.

Complex rotations ($M = 22.6$ s) were significantly slower than planar ones ($M = 11.7$ s): $F_{1,11} = 66.7$, $p < .001$. But again, there was no evidence that any of the techniques was particularly good or bad for complex manipulations (technique × rotation interaction, $F_{2,20} = 1.2$, $p = .33$).

Only tasks that were completed within the 40 s time limit were included in this analysis. To check that these results were not adversely influenced by different rates of incomplete trials in different conditions, we analysed the number of incomplete trials using the same $3\times3\times3$ ANOVA. This analysis further supports the results above. Timed-out tasks were significantly more prevalent when using fewer points of contact ($F_{1,16} = 7.3$, $p = .01$), with means of 1.3, 0.6 and 0.2 timeouts per condition with one-, two- and three-touch respectively. There were significant effects for difficulty ($F_{1,14} = 9.4$, $p < .01$) and rotation ($F_{1,11} = 14.8$, $p < .01$); but there was additionally a significant *technique × rotation* interaction ($F_{1,16} = 7.9$, $p < .01$), due to a much more dramatic increase in timed-out tasks between planar and complex tasks with one-touch than with two- and three-touch (Figure 9). As before, the technique × difficulty interaction was not significant ($F_{3,31} = 0.8$, $p = .15$).

Characterising Interaction with the Techniques
The analysis above shows that the participants completed tasks more quickly when given more points of contact for interaction, and that the benefits of doing so become larger in more complex tasks. In order to better understand the strengths and weaknesses of each of the techniques for particular types of object manipulations, we now further scrutinise data on the time spent conducting particular types of manipulations, and the object regions used to do so.

To conduct this analysis we broke down the TCTs into time spent performing translation, planar rotation and spatial rotation. For the one-touch technique, this can be separated by time spent touching each dedicated area on the objects. For the two-touch technique, it is done by separating time spent inside and outside the translation-only area, and by measuring time spent using the second finger. For the three-touch technique, it is separated into time spent touching with each finger. Note that the sum of all movement types can be more than the TCT for the two- and three-touch techniques, since the user can perform multiple movements at the same time.
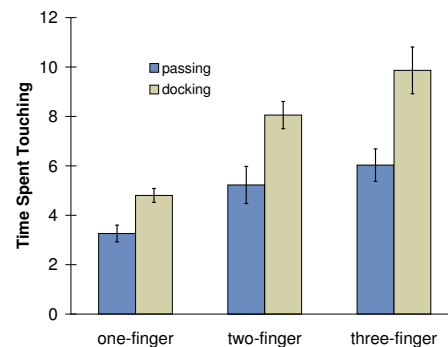

Figure 10. Mean time spent touching the object, separated into translations, planar rotations and spatial rotations.

We analysed the decomposed TCTs using a 3×3 within-subjects ANOVA for factors technique (one-, two-, three-touch) and movement type (translation, planar rotation, and spatial rotation).

There was a main effect of technique ($F_{2,18} = 8.5, p < .01$). Figure 10 shows mean time spent for each technique. Post-hoc comparisons show that participants touched the screen significantly less with the one-touch technique than with the three-touch technique ($p < .001$) and marginally less than with the two-touch technique ($p = .06$). The difference between the two-touch and three-touch techniques was not significant ($p = .12$). This effect is in direct contrast to the main effect of technique for TCTs alone. This contrast suggests that participants spent more time performing cognitive processing than interaction with less DOF and that this resulted in higher TCTs. Experimenter observations also confirmed that participants tended to have more difficulty with mental rotations when using the one-touch technique. Note, however that the measures fail to discriminate between manipulations that occur in parallel and in series, so this result should be cautiously appraised.

There was a significant interaction between technique and movement type ($F_{2,23} = 18.7, p < .001$) shown in Figure 11. Post-hoc comparisons show that for one-touch interaction, participants spent significantly more time performing spatial rotations than either translations ($p < .001$) or planar rotations ($p < .001$) and that for three-touch interaction, participants spent significantly more time performing translations than either planar rotations ($p < .001$) or spatial rotations ($p < .01$). All other pairwise differences were not significant ($p > .05$). This interaction shows that participants typically spent an approximately equal amount of time performing rotations with all three techniques. Furthermore, the larger amount of translations in the three-touch condition may be because participants were able perform translations in tandem with the other types of rotation. This result illus-
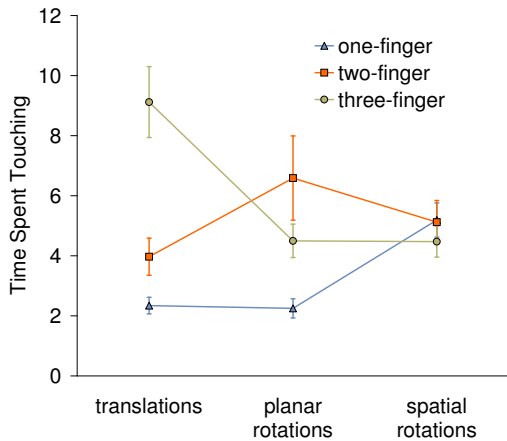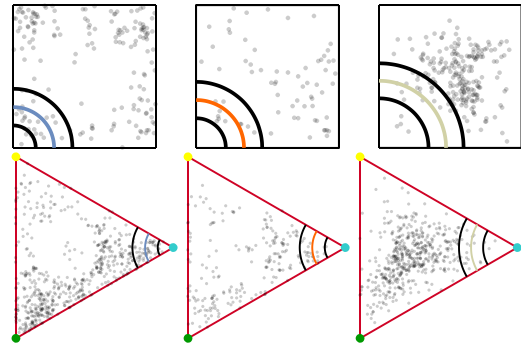
Figure 12. Touch locations on a typical face of the cubes in the passing task (top) and tetrahedrons in the docking task (bottom) separated into one-touch (left), two-touch (middle), and three-touch techniques (right). The coloured arcs represent the mean distances to the nearest corner for each touch location, black arcs represent the standard deviation from these means.

trates very well that simultaneity of movements provides a strong advantage for multiple DOF interaction.

The results for the docking task are sufficiently similar to not warrant reiteration here. However, the fact that both tasks have the same main effect, interaction and pairwise differences further strengthens this result.
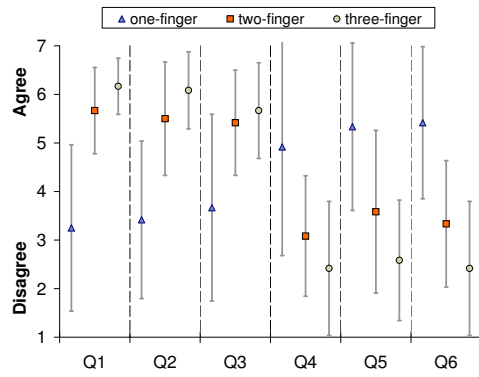
Touches
We observed during the experiment that participants tended to use object corners for spatial rotations much more with some techniques. We recorded the locations of every touch intended for spatial rotation made by each participant and rendered each point using a constant transparency. Patterns clearly show that for the one-touch technique, users concentrated their touches at the corners and for the two- and three-touch techniques, the touch locations were more central. Figure 12 shows a typical face of both the cube from the passing task and the tetrahedron from the docking task for each technique. We also recorded the number of times the participants missed the objects completely and found that this occurred most frequently with the one-touch technique.

Subjective Ratings
Figure 13 shows the average scores on the follow-up questionnaire. For the docking task, 9 participants preferred the three-touch technique and 3 preferred the two-touch technique. For the passing task, 7 participants preferred the three-touch technique, four preferred the two-touch, and 1 preferred the one-touch technique. Overall, 7 participants preferred the three-touch, 3 preferred the two-touch, and 1 claimed there was no clear winner.

All subjective data shows a clear order of preference from three-touch (best), two-touch, to one-touch (worst). Participants consistently rated the three-touch technique as the easiest to use (Q1) with the most appropriate reaction (Q2), as the least difficult to control (Q4) and rotate – both in the plane (Q5) and spatially (Q6). Also, the three-touch technique was most preferred for docking, passing and overall. The two-touch technique was rated second in all categories and the one-touch third, though with much higher variance.

Figure 11. There was a significant interaction between technique and movement type. For the one-touch technique, participants performed more spatial rotations than translations or planar rotations and for the three-touch technique, participants spent more time performing translations than planar or spatial rotations. The difference in movement type did not differ significantly for the two-touch technique.

Q1	I found the technique easy to use
Q2	With this technique, the object reacted as I expected it to
Q3	I could easily move an object to where I wanted
Q4	I found the technique difficult to control
Q5	I found it difficult to turn objects in the plane
Q6	I found it difficult to roll objects over

Figure 13. Mean ratings and (standard deviations) on the follow-up questionnaire. Users rated their level of agreement on a scale from 1 (strongly disagree) to 7 (strongly agree).

Overall Discussion

Our study showed that the techniques that use a higher number of touches were better both in terms of performance and user preference. These benefits likely appeared because the higher number of touches provided users with the opportunity to independently control more degrees of freedom. This type of freedom provides increased flexibility for how users decide to perform the interactions.

Our study showed that one-touch interaction was rated as difficult to use and resulted in the worst performance. This result implies that one-touch interaction (as designed here) was not efficient for interacting in shallow-depth 3D on the table. One response would be to redesign the one-touch interaction technique (one alternative is discussed below). Perhaps a more important consequence is that most existing hardware input technology is currently insufficient for supporting multi-touch interaction. Our results suggest that *multiple independent inputs for each person* at the table will be beneficial for both performance and satisfaction.

One concern we had when initially developing the techniques was that the complexity of multi-touch interactions would prove confusing and deter users from its acceptance. In contrast, allowing users separate and simultaneous control of rotation and translation provided a more preferred interaction with better performance. From watching people use these techniques, one could see that their interactions became more natural and easy as the number of touch points increased. Users are not only capable of this more engaged, complex control, but prefer it.

Generally, participants in our study were both intrigued and excited by all three techniques. This enthusiasm is likely due to the novel ability to use digital objects in a way that was more similar to their experiences with physical objects on tables. Users commented that with these techniques it felt "more like I was touching it" and that "I almost want

to look" under the objects. However considering the actual TCTs in comparison with what people are capable of with physical objects, there is considerable scope for future research refining these and other new techniques for manipulating shallow-depth 3D objects. Nonetheless, these techniques do provide the first steps toward enabling the more complex 3D interactions with which we are familiar.

ALTERNATIVE TECHNIQUES

In light of the results of our study, we have explored alternative designs for our interaction techniques. Specifically, we believe that a redesign of the one-touch technique might make for a feasible method for interacting on tables incapable of multi-user, multi-touch, direct-touch interaction. Furthermore, our multi-touch techniques typically assign object transformations based on the movement of *every* finger. Another way of implementing bimanual, multi-touch rotation would be to use the additional touches to introduce constraints that limit chosen aspects of the interaction. Such interactions have been shown to be an approach that users can easily cope with, due to the kinesthetic feedback [12].

Alternative One-Touch Technique

The results of our experiment showed that, while spatial rotation interactions were accessible from both edges and corners, people typically made almost exclusive use of the corners. We also found that users had difficulty acquiring the corners and would frequently miss the object entirely. In our new design, the 3D rotation previously available on the entire surface of the object is only allowed at the corners and the user may acquire the corner by selecting anywhere inside a sphere about each vertex of the polygon. The object still has a translate-only region in the centre of each face, but the remaining parts of the object allow only planar RNT interaction. This new technique still benefits from the property that the selected point remains under the user's finger.

Alternative Multi-Touch Technique

One of the disadvantages of both multi-touch techniques used in our study is that the point of contact may not remain under the user's finger once a rotation is performed with the finger on the non-dominant hand. We propose an alternative three-touch technique that constrains the effect of the primary finger based on the presence or absence of contact of the thumb and/or the finger on the non-dominant hand. When the user manipulates the object with their primary finger and no other finger is touching, the object reacts as it would in the one-touch technique. When the user uses both the thumb and the index finger, planar rotation is performed as in the three-touch technique. The user can then limit the movement to translation-only by touching the table with a finger on the non-dominant hand. This technique also has the advantage that the point of contact remains under the user's finger. It also corresponds to the way physical objects react, in that additional points of contact allow for more precise, constrained motion.

CONCLUSION & FUTURE WORK

In this paper, we have introduced and studied three new rotation and translation techniques for interacting in shallow-

depth 3D on a digital table. In light of insights gained through our study we suggest two additional approaches to providing this type of interaction. These techniques are the first steps toward realizing our vision of shallow-depth 3D interactions in the digital realm which are much more closely aligned to those we are familiar with on traditional tables.

We provide guidelines for the design of direct-touch interaction in 3D and a user study that reinforces these guidelines. Our study also shows that the greater expressive power of more touch points can improve performance and user satisfaction for direct-touch manipulation of 3D objects on a digital table. In detail:

- Shallow-depth was easily understood and interpreted as a natural environment.

- People are generally enthusiastic about manipulating 3D objects on digital tables.

- A higher number of touches allows more natural and flexible interaction.

- One-touch interaction should be re-explored.

- Multiple points of identifiable touches should be supported in hardware.

- People are not only capable of separable simultaneous control of rotation and translation, but prefer it.

While continuing to explore further refinements of our techniques, we also intend to empirically explore what type of feedback is appropriate in a shallow-depth 3D interface. We are specifically interested in addressing issues of shading and parallax that are most closely associated with the use of 3D in a collaborative setting. Our vision is an interface where multiple users can interact simultaneously from any side of the table combined with the rich expressive interactions available in 3D.

REFERENCES
1. Agarawala, A. and Balakrishnan, R. Keepin' it real: Pushing the desktop metaphor with physics, piles and the pen. In *Proc. CHI*, pp. 1283–1292. ACM Press, 2006.

2. Deitz, P. and Leigh, D. DiamondTouch: a multi-user touch technology. In *Proc. UIST 2001*, pp. 219–226. ACM Press, 2001.

3. Dragicevic, P. Combining crossing-based and paper-based interaction paradigms for dragging and dropping between overlapping windows. In *Proc. UIST*, pp. 193–196. ACM Press, 2004.

4. Fröhlich, B., Hochstrate, J., Skuk, V., and Huckauf, A. The GlobeFish and the GlobeMouse: Two new six degree of freedom input devices for graphics applications. In *Proc. CHI*, pp. 191–199. ACM Press, 2006.

5. Han, J. Y. Low-cost multi-touch sensing through frustrated total internal reflection. In *Proc. UIST*, pp. 115–118, New York, NY, USA, 2005. ACM Press.

6. Hancock, M. S., Vernier, F., Wigdor, D., Carpendale, S., and Shen, C. Rotation and translation mechanisms for tabletop interaction. In *Proc. Tabletop*, pp. 79–86. IEEE Press, 2006.

7. Jacob, R. J. K., Sibert, L. E., McFarlane, D. C., and M. Preston Mullen, J. Integrality and separability of input devices. *TOCHI*, 1(1):3–26, 1994.

8. Kruger, R., Carpendale, S., Scott, S. D., and Greenberg, S. Roles of orientation in tabletop collaboration: Comprehension, coordination and communication. *JCSCW*, 13(5-6):501–537, 2004.

9. Kruger, R., Carpendale, S., Scott, S. D., and Tang, A. Fluid integration of rotation and translation. In *Proc. CHI*, pp. 601–610. ACM Press, 2005.

10. Matsushita, M., Iida, M., Ohguro, T., Shirai, Y., Kakehi, Y., and Naemura, T. Lumisight table: a face-to-face collaboration support system that optimizes direction of projected information to each stakeholder. In *Proc. CSCW*, pp. 274–283. ACM Press, 2004.

11. Nakashima, K., Machida, T., Kiyokawa, K., and Takemura, H. A 2D-3D integrated environment for cooperative work. In *Proc. VRST*, pp. 16–22. ACM Press, 2005.

12. Sellen, A. J., Kurtenbach, G. P., and Buxton, W. S. The prevention of mode errors through sensory feedback. *HCI*, 7(2):141–164, 1992.

13. Shen, C., Vernier, F. D., Forlines, C., and Ringel, M. Diamondspin: an extensible toolkit for around-the-table interaction. In *Proc. CHI*, pp. 167–174, New York, NY, USA, 2004. ACM Press.

14. Ståhl, O., Wallberg, A., Söderberg, J., Humble, J., Fahlén, L. E., Bullock, A., and Lundberg, J. Information exploration using the pond. In *Proc. CVE*, pp. 72–79. ACM Press, 2002.

15. Wang, Y., MacKenzie, C. L., Summers, V. A., and Booth, K. S. The structure of object transportation and orientation in human-computer interaction. In *Proc. CHI*, pp. 312–319. ACM Press, 1998.

16. Ware, C. *Information Visualization: Perception for Design*. Morgan Kaufmann, 2nd edition, 2004.

17. Zhai, S. and Milgram, P. Quantifying coordination in multiple DOF movement and its application to evaluating 6 DOF input devices. In *Proc. CHI*, pp. 320–327. ACM Press, 1998.