2D Similarity Transformations on Multi-Touch Surfaces

Behrooz Ashtiani, Wolfgang Stuerzlinger

York University, Toronto, Canada

ABSTRACT

We present and comparatively evaluate two new object transformation techniques for multi-touch surfaces. Specifying complete two-dimensional similarity transformations requires a minimum of four degrees of freedom: two for position, one for rotation, and another for scaling. Many existing techniques for object transformation are designed to function with traditional input devices such as mice, single-touch surfaces, or stylus pens. The challenge is to map controls appropriately for each of these devices. A few multi-touch techniques have been proposed in the past, but no comprehensive evaluation has been presented.

XNT is a new three-finger object transformation technique, designed for multi-touch surfaces. It provides a natural interface for two-dimensional manipulation. *XNT* and several existing techniques were evaluated in a user study. The results show that *XNT* is superior for all tasks that involve scaling and competitive for tasks that involve only rotation and positioning.

Keywords: multi-touch, evaluation, 2D transformation, similarity transformation, rotation, translation, scale, touch techniques, direct manipulation.

Index Terms: H5.2 [Information Interfaces and Presentation]: User Interfaces - Interaction Styles.

1 INTRODUCTION

Multi-touch surfaces have grown in popularity in recent years, partially due to the efforts of major technology companies. The multi-user collaborative experience of Microsoft's *Surface* [1], SMART Technologies' *SMART Table* [2], Circle Twelve's DiamondTouch [3], and Perceptive Pixel's *Interactive Media Wall* [4] is designed to facilitate a more natural and intuitive form of human-computer interaction.

Digital tabletops enable face-to-face collaboration to multiple users, allowing them to gather around a table and observe the displayed content from multiple viewpoints around the table. These multiple viewpoints introduce a design challenge: visual content properly oriented and legible to one user may not be so to another [21]. Inevitably, this necessitates techniques that enable users to freely transform visual content to match their viewpoints. Scaling is important as well, as users may need to access more or less detail. Hence, interactions with the position, size, and orientation of the displayed objects are necessary. The combination of these three transformations is defined by similarity transformations, a generalization of rigid body transformations.

4700 Keele Street, Toronto, ON, Canada. abehrooz@cse.yorku.ca

2 LITERATURE REVIEW

Various techniques have been developed for similarity transformations of visual content on touch surfaces. In general, such transformations are specified by three components: translation, rotation, and uniform scale. An obstacle in designing transformation techniques is the number of degrees of freedom (DOF) provided by the input. E.g., a mouse offers direct control of only two DOF. Multi-touch devices are inherently capable of producing many more DOF for input. This invites the design of techniques that are able to encapsulate all components of similarity transformations into a single interaction gesture.

2.1 Techniques based on 1 DOF Input

Some transformation techniques do not need to utilize all the available DOF to perform a task. For example, menus or controls for explicit specification of object parameters require only 1 DOF of input. E.g. in Microsoft© *Paint* rotations are performed with a menu. Autodesk© *AutoCad* [5] allows all parameters including position of the objects to be specified explicitly.

For tabletops, Wu and Balakrishnan's *Parameter Adjustment Widget* [43] is a two-finger technique where one finger selects the object and the other finger uses a menu with up and down arrows to change the rotation angle. This technique gives the user very precise control, at the expense of fluidity and naturalness of interaction, and suffers from unnecessary cognitive overhead [20].

2.2 Techniques based on 2 DOF input

Techniques that take advantage of the two DOFs afforded by the input device employ different strategies to produce the translation and rotation components of object transformations. One approach involves separating the translation and rotation components and manipulating each independently, for example with a mode switch or different handles. Adobe© *Photoshop* and Microsoft© *PowerPoint* are examples of applications that employ such "handle" controls for rotating and resizing objects. These handles are usually placed at the edges and corners of a selected object. Depending on the handle, only one component of transformation (e.g. rotating, resizing, moving) is manipulated at a time. For digital tabletops, a technique known as "corner to rotate" has been presented. It enables translation when the area inside the object is touched and dragged. Dragging on the corners of the object causes rotation around the center of the object [20,33,36,37,39].

To alleviate performance issues when manipulating each component of transformation independently [16], new techniques able to perform composite translation and rotation emerged. *DiamondSpin* [12,36] allowed automatic orientation of documents while they are being translated, by orienting the top of the document towards the center of the digital table. Other implementations of automatic orientation include *InfoTable* [32], *ConnecTable* [39], and *Magnetic Poetry* [34]. A clear limitation of automatic orientation methods is that they limit the number of position-angle combinations that can be easily achieved [15].

Kruger *et al.* [19] report that automatic orientation of an object, based on the location of its manipulator, or the territory the object is in, does not always yield the intended results. Overcoming the innate limitations of “corner to rotate” and automatic orientation techniques necessitates that any new technique not only allows full control over all components of transformation, but to do so in a single integrated gesture. The latter requirement is based on studies [15,16,20], which suggest that the human mind perceives translation and rotation as integral and inseparable.

The window rotation technique by Beaudouin-Lafon [7], *Drag* [25], and *Rotate’N Translate* (*RNT*) [20] are mathematically equivalent integral techniques based on “simulated friction”. There, translating an object causes it to simultaneously rotate, according to the point of contact and the direction of movement. These techniques simulate the physics of a real piece of paper on a table with a single finger. Then, the location of the touch influences the rotation behaviours when moving the page. Figure 1 visualizes this. Integral techniques overcome the limitation of automatic orientation while outperforming “corner to rotate” approaches [20].

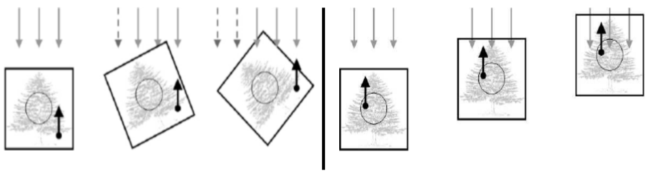

Figure 1: Simulated friction in *RNT* [20]: dragging on the exterior region results in simultaneous translation and rotation (left), dragging on the central region results in translation (right).

2.3 Techniques based on Higher DOF input

Lui *et al.* describe a set of techniques collectively known as *TNT* [21]. Touch, hand motion, and the angle of rotation are detected using a Polhemus *Fastrak* tracking system (Figure 2). In one approach, *TNT-hand*, the tracker is placed inside a “finger sleeve”, worn on the index finger. In another approach, called *TNT-block,* the tracker is placed inside a cylindrical block, which can be held and rolled between the fingers. A direct pre-cursor to the work on *TNT* is *Bricks* [11]. The results of their user study indicate that *TNT-block* and *TNT-hand* techniques outperform the *Rotate’N Translate* (*RNT*) technique.

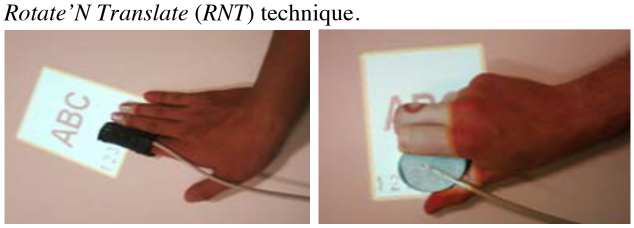

Figure 2: TNT-hand (left) and TNT-block (right) techniques [21] use embedded trackers to simultaneously translate and rotate digital objects.

Matsushita *et al.* [23] describe a dual touch technique called *Bistroke* for rotating a displayed map on PDAs. Placing a finger on the screen sets the center of rotation and stroking the screen with the stylus pen rotates the map by the change of angle between contact points. Wu *et al.* [43] discuss a similar technique for digital touch tables, with the additional feature that permits lifting the original pivot finger once the rotation is confirmed.

Sticky Fingers [22] is a two-finger technique that supports scaling, translations, and rotations. When touching the object, it “sticks” to the fingers and moves with them as they move; the distance between the touched points and their relative orientation change the scale and rotation respectively.

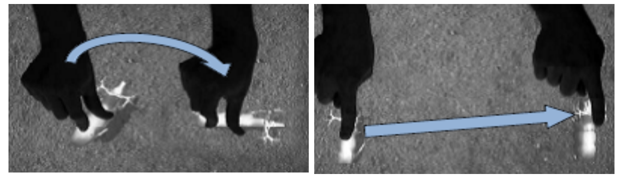

Figure 3: *Sticky Fingers* technique [22]: translation (left), simultaneous translation and rotation (right).

3 MOTIVATION AND DESIGN

Multi-touch interfaces enable users to perform two-dimensional transformations, resembling natural interaction with the physical world, except scaling. A multi-touch technique is henceforth referred to as full-2D, if it supports all components of similarity transformations, i.e. translations, rotation, and scaling.

There have been studies on direct and indirect mapping of multi-touch input [18,35], separability and integrality of transformations [16,28,40], and uni-manual vs. bi-manual interaction [13,18,27]. Yet and to our knowledge, no comparative study on the performance and accuracy of full-2D techniques exists in the literature. In fact, even for partial 2D manipulation techniques, there is only a single study [21] and it does not use a multi-touch solution. In this paper, based on guidelines available in the literature and a pilot study performed in our lab, we introduce two new 2D techniques and compare their performance and accuracy to *TNT-hand* [21] and *Sticky Fingers* [22].

TNT-hand was chosen for multiple reasons. It has been comparatively evaluated with other techniques, such as *RNT* [19], and was shown to be superior. Also, while *TNT-hand* does not support scaling and was originally implemented using a tracker, it is possible to emulate this technique with a touch system. To avoid confusion with the original implementation, our version will be called *TNT-Touch* hereafter. The *Sticky Fingers* technique is a commonly used full-2D transformation technique and serves as a baseline condition. We closely follow the same experimental methodology as the *TNT* study [21] in an effort to bridge the gap between the two studies.

3.1 Direct vs. Indirect Touch

Since the focus of this study is on performance and accuracy of techniques on multi-touch surfaces, we opted to compare them in the direct-touch condition. Kin *et al.* [18] report an 83% reduction in selection time when direct-touch is used. This confirms several earlier observations, which compared direct single-touch to indirect mouse interaction [17,30]. Schmidt *et al.* [35] demonstrate that for a multi-touch translation and scaling task, direct-touch performs substantially faster than indirect touch. Reisman *et al.* [31] justify the design of their direct-touch 3D technique citing popularity and intuitiveness of previous 2D direct touch techniques.

3.2 Intuitive Precise Control

Our focus here is similarity object transformations. For this task direct manipulation allows objects to be controlled more intuitively than gesture-based manipulation [31]. While there are gesture-based techniques [6,10,24,26,42], these are usually targeted at higher-level functionality, such as copy & paste, and grouping. Also, designing interaction techniques based on physical world analogies can often be restrictive and limiting [27]. This is particularly important with direct touch interaction.

In touch interfaces, objects may be partially or completely occluded by the fingers or hand(s). Hence, direct-touch interaction techniques need to tackle the *fat finger problem* [8,41], yet still enable precise and intuitive object selection.

3.3 Uni-manual vs. Bi-manual

In designing interaction techniques for their *Room Planner* application, Wu and Balakrishnan [43] argue that their multi-finger uni-manual technique is an improvement over a similar bi-manual technique proposed by Bier *et al.* [9]. This advantage is attributed to the closer link between fingers of the same hand in the kinematic motor control chain [13]. Moscovich and Hughes [27] discuss the higher compatibility of uni-manual techniques with two-dimensional transformation tasks compared to bi-manual techniques. For the integral task of simultaneous translation, rotation, and scaling of objects, uni-manual techniques perform better. Nacenta *et al.* [28] report the common use of uni-manual operation when the object is small, or if the second hand is otherwise occupied. In a multi-user setting, identifying users becomes easier with a uni-manual technique, since the fingers of a given user's hand remain in close proximity to each other.

3.4 Hand Posture and Number of Fingers

Per se, two fingers provide sufficient input for performing all full-2D transformation tasks. Hence, many existing techniques, such as *Sticky Fingers* [22], rely only on two fingers. The design of *TNT-hand* [20] was based on an observational study of how real sheets of paper are rotated and transported on tables. Lui *et al.* [20] reported popular use of an "open palm transfer"; where fingers are extended and object is moved with the palm or fingers.

We expected "open palm transfer" not to work well on a multi-touch system as placing the palm of the hand on the surface restricts hand movement due to increased friction. Reducing the contact to the fingertips improves this but may cause loss of stability and control. In a pilot observational study and to investigate the effect of friction, we asked four participants to first move and rotate a sheet of paper on the tabletop screen with only their fingertips using different combination of fingers and then to imitate the same movements directly on the tabletop. When using two fingers, all other fingers have to be held up, which makes this posture more fatiguing. Conversely, using all the fingers provided the most comfort at the expense of more friction. Using the triad of thumb, index, and middle fingers together improves stability and control while keeping friction at a reasonable level.

3.5 Integral vs. Separated

In their study of input devices, Jacob *et al.* [16] discussed the importance of matching the control structure of the technique to the perceptual structure of the task. Research by Wang *et al.* [40] shows also a parallel interdependency between translation and rotation controls. They conclude: "Constraints or interruption on the integration of object manipulation may result in a structural inefficiency." They further find that the total task completion time may increase significantly once the parallel structure of translation and rotation is replaced with a serial structure. Reisman *et al.* [31] argue that continuous contact in integral methods makes them convenient and easy to use. In our study we compare an integral and a separated technique directly to investigate this issue.

3.6 The Design of XNT and XNT-S

We present a list of design requirements and guidelines extracted from previous studies and our own pilot for improved full-2D transformation techniques:

- Support direct, uni-manual, multi-finger interaction
- Provide intuitive, yet precise control
- Based on a stable and comfortable hand posture
- Able to perform integral object transformations

We hypothesized that *TNT-Touch* will perform well as it follows these guidelines closely. However, it only supports rigid object transformations (i.e., no uniform scaling). *Sticky Fingers* does not support a stable and comfortable hand posture. It also suffers from the *fat finger problem* during selection and manipulation of small objects. Hence, we expected *Sticky Tools* to perform poorly in tasks involving scaling.

Our new technique, *XNT*, adheres to all these guidelines. It permits direct uni-manual interaction with three (or more) fingers. Object selection is based on the first finger that touches or, if no object was hit, whatever object is first intersected by the centroid of all touch points. As long as no object is intersected, the user can use this centroid to "search" for an object to manipulate, which is crucial for enabling selection of very small objects, even down to the size of a single pixel. Benko *et al.* [8] described a similar dual-finger midpoint selection technique. In our pilot study, we observed a notable user preference for three-finger interaction with the thumb, middle, and pointing fingers, especially for rotation tasks. *XNT* is integral, thus it allows simultaneous manipulation of translation, rotation, and scale. However, to investigate the role of integrality and separation of multi-touch input in two-dimensional object transformation tasks, a separated version called *XNT-S* was also implemented. *XNT-S* assigns the components of similarity transformation to touch input so that each component has to be manipulated independently. The next section discusses implementation details for our techniques and *TNT-Touch* and *Sticky Fingers*.

4 Implementation

For a virtual object to be manipulated, it first needs to be selected. Conversely, an object is unselected once there are no fingers touching. Also, transformation of objects requires a reference point around which translation, rotation, and scaling are applied. Using the object center as the reference point creates a potential disconnect between interaction and visual feedback. For instance, touching a rectangular object at a corner and performing a rotation can cause the rectangle to escape from under the user's fingers if the rotation is applied at the center. A better strategy is to assign the first contact point with the object, i.e., the first point the user touches, as the reference point. When multiple fingers are manipulating the object, the reference point can be the centroid of these points. However, the assignment of the reference point is closely tied to the selection mechanism for each technique.

To describe the multi-touch aspects of techniques in a precise, yet compact manner we propose a notation that lists the number of fingers followed by letters that signify the transformations afforded by that number of touches. For transformations "T" refers to translation, "R" to rotation, and "S" to scaling. Hence, "1T 2TR 3TRS" denotes a technique that enables translation with one finger; simultaneous translation and rotation with two; and simultaneous translation, rotation, and scale with three.

4.1 Sticky Fingers

This technique allows selection of the objects with one or two fingers. Selection with single touch occurs when the touch point lands on the object and that serves as the reference point. Selection with two fingers occurs when both fingers are on the object, and the reference point is the centroid. We classify this technique as "1T 2TRS", as it is able to perform full-2D transformation with 2 fingers. Both fingers must be on the object for full-2D transformations, as described previously [22] and a single touch on the object will only translate.

4.2 XNT

This technique allows selection using any number of fingers. The first finger to touch the surface selects the object. Further, an object can be selected even if none of the fingers are in contact with the object, but if the centroid of the touched points gets into contact with an object. The reference point is set to the selection point, i.e. the first contact point or the centroid, and remains until that point is no longer valid. For instance if the finger that made the first contact with the object is lifted from the surface, the algorithm looks for a suitable replacement selection point with priority given first to the fingers and then to the centroid. This technique is classified as "1T 2TRS 3TRS" and performs full-2D transformations on the selected object. In other words, when two or more fingers are present on the surface, and if any of the fingers or the centroid is in contact with the object, full-2D transformation is enabled.

4.3 XNT-S

This technique shares its selection mechanism and reference point assignment with *XNT*. It differs in how it performs the transformations, in that *XNT-S* can be described as "1T 2S 3R". Hence, the number of fingers directly defines the transformation that is applied. This separation may result in higher accuracy since complex transformations happen in series rather than parallel.

4.4 TNT-Touch

This technique is a touch-based reimplementation of *TNT-hand* [21]. It uses the index and the middle fingers exclusively and requires them to be both present and held close together. This was achieved in our study by wrapping these two fingers together using an adjustable strap. For selection both fingers are placed on the object. Their centroid then becomes the reference point. This technique can be classified as "2TR" and is used as a point of comparison with previous research.

5 Methodology

This section details the experimental design and procedures in our comparative study of two-dimensional similarity object transformation techniques.

5.1 Participants

Twelve volunteer participants were recruited, aged 19 to 30 years (*mean* = 24, *SD* = 4). Nine were male. On average, participants report 7 hours of computer use per day, 4 hours of gaming per week, and 12 hours of mobile touch device usage per week.

5.2 Apparatus

The experiment was performed on the *Multi-User Laser Table Interface* (*MULTI*) platform [38]. With multiple 120 Hz NaturalPoint OptiTrack *FLEX:C120* optical cameras located underneath the table surface, the interface captures and reports the movement of bright points of light, normally generated via laser pointers. This tabletop uses highly off-axis NEC WT600 projectors. Hence, it uses a very diffuse projection surface, with a diffuse coating applied to one side. This impedes the use of standard touch-detection technologies, such as FTIR [14] or diffuses illumination. To add touch-detection capabilities to this system, we constructed multiple "fingerlings", i.e. gloves for only a single finger, where a red LED is placed at the tip of the finger (Figure 4). These fingerlings were worn on the thumb, index and middle finger. Due to the diffuse surface and with an appropriate brightness threshold, the system then detects events only when the LED's in the fingerlings touch (or are very close to) the surface.

One of the reasons this setup was chosen for this study was its high touch detection rate of 120 Hz, equivalent to the sampling rate of a mouse. Also, the tabletop has 1536x1024 pixels, which is more than other systems based on a single projector.

The fingerlings make accurate detection of touch very easy as only bright spot detection is required. Hence, touch detection is extremely reliable. Another potential major advantage of the fingerlings is that the LEDs can be blinked in appropriate patterns synchronized with the camera refresh rate. This enables reliable distinction between fingers and due to the high refresh rate, these patterns are effectively invisible [29]. This idea can be naturally extended to multiple hands and/or users. However, we did not use this option in the current user study.

Figure 4: Fingerlings designed to add multi-touch capability (left), reaching to transform a 2D virtual object (right).

5.3 Tasks

The main user study involves three precise docking tasks, modelled after Liu *et al*. [21] in their comparison study of *TNT* and *RNT*. They have been modified to test also the resizing of objects. All tasks involve matching a rectangular object, initially sized 92×132 pixels, to the position, orientation, and size of a transparent rectangular target. There are four possible locations where the object and the target can appear, as depicted in Figure 5. A single trial is defined by a starting position where the object initially appears and a target position. As in previous work, only a subset of all possible paths between the 4 targets was tested. The direction for each of the 4 tested paths was randomly chosen for a given trial. However, as many different trials are performed for each task, the study explored all 2×4=8 paths systematically. The order of the paths was also randomized.

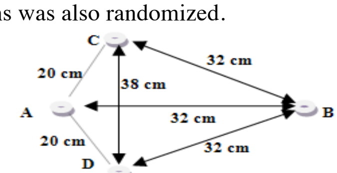

Figure 5: The four possible initial and target positions. Arrows indicate the direction of all tested paths.

The physical size of the object on the screen was 11x16 cm. The object is always placed in an upright position at the starting position with zero degree rotation and unit scale. Targets can assume five different orientations: 0° (no rotation), 45° clockwise, 135° clockwise, 45° counter-clockwise, and 135° counter-clockwise. There are three possible target sizes: unchanged (scale = 1), half (scale = 0.5), and quarter (scale = 0.25).

In the *Translate and Scale Task*, targets are always in the upright orientation, but vary in scale. For each possible path, both half and quarter sizes are tested. This yields 4 paths × 2 sizes = 8 trials. In the *Translate and Rotate Task* the size of the targets always remains unchanged. All four non-zero orientations were tested. There are 4 paths × 4 orientations = 16 trials in this task. In the *Translate, Rotate, and Scale Task*, all orientations and sizes are tested, including the 0° orientation and unchanged size. Hence, this task involved 4 paths × 5 orientations × 3 sizes = 60 trials.

5.4 Procedure

Prior to collecting data, the experimenter instructed the participants on using the fingerlings, and explained the tasks briefly. A tutorial session was given before each new task to explain it and the techniques. Participants were given about one minute to get accustomed with techniques and tasks. Tutorial sessions ended once participants voiced their readiness for starting the next set of trials. None of the participants required more than a few trial runs to become familiar with a given task. Participants were also given an optional break at that time.

A single trial began as soon as the object and target appeared on the screen in line with previous research [30]. Trials ended once the object was within a preset error threshold for position, orientation, and size, when the participant lifted their finger off the surface. For this, the object's center had to be within 10 pixels of the target's center, its orientation within 5 degrees of the orientation of the target, and its size within 5% of the target size. These values were selected based on pilots on error thresholds for which the inconsistencies in position, orientation, and size between the object and the target were visually apparent.

After the experiment, participants completed a questionnaire, with questions on demographic information and their perception of performance, accuracy, and fatigue of each technique.

5.5 Experimental Design

Each of the 3 tasks used a different repeated-measures factorial design depending on the number of testable factors:

- Translate and Scale: 3 techniques × 8 paths × 2 sizes
- Translate and Rotate: 4 techniques × 8 paths × 4 orientations
- Translate, Rotate and Scale: 3 techniques × 8 paths × 4 orientations × 3 sizes

Since *TNT-Touch* was only capable of performing the translation and rotation task only 16 trials per participant were recorded. All other techniques were capable of performing all tasks, which yielded data for 8+16+60=84 trials per participant. The order of techniques was assigned by a balanced Latin square in each case. Overall, data for 3216 trials was collected: 12 participants × (3 techniques × 84 trials + 1 technique × 16 trials).

6 RESULTS

Results in this section are organized based on task.

6.1 Translate and Scale Task

A repeated-measure ANOVA performed on the data revealed a significant main effect of technique on task completion time ($F_{2,11}$= 4.34, $p < .05$). A follow-up Tukey-Kramer test revealed a pair-wise difference between the *XNT* and *Sticky Fingers* techniques. In the Translation and Scale task, *Sticky Fingers* and *XNT-S* performed 84% and 11% slower than *XNT*. See Figure 6.

There was no significant effect of technique on position error ($F_{2,11} = 1.64, p > .05$), but technique had a significant effect on orientation error ($F_{2,11} = 20.84, p < .0001$). A post-hoc Tukey-Kramer test placed *XNT-S* in a separate group due to the small average error.

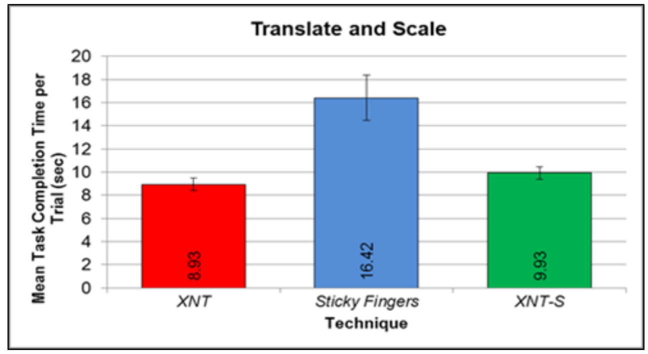

Figure 6: Mean task completion time by technique for the Translate and Scale task. Error bars show ±1 SE.

Also, there was a significant effect of technique on size error ($F_{2,11} = 4.25, p < .05$), and Tukey-Kramer revealed that both *XNT* and *Sticky Fingers* are distinct from each other, but not from *XNT-S*. Figure 7 visualizes the mean errors.

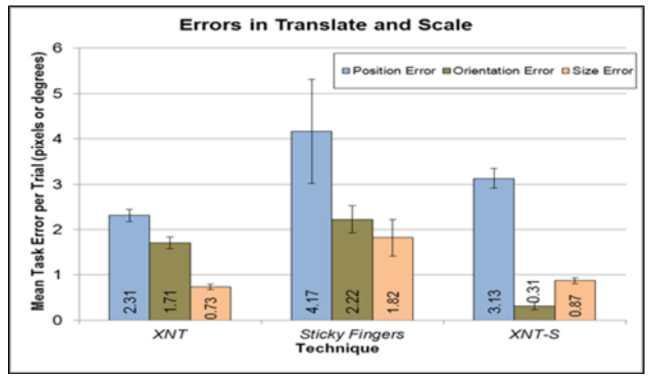

Figure 7: Mean task Error by technique for the Translate and Scale task. Position and Size errors are reported in pixels and orientation errors in degrees. Error bars show ±1 SE.

We found a significant interaction between technique and target size ($F_{2,11} = 4.81, p < .05$). The Tukey-Kramer test indicated *Sticky Fingers* to be in a separate group from the other two. While decreasing target sizes increased task completion time across all techniques, *Sticky Fingers* experienced the most performance reduction. *Sticky Fingers* was 135% slower when the target size was changed from half to quarter. In comparison *XNT* and *XNT-S* became only 63% and 12% slower respectively at quarter size.

6.2 Translate and Rotate Task

A repeated-measure ANOVA revealed a significant main effect of technique on task completion time ($F_{3,11}$ = 11.92, $p < .0001$). A Tukey-Kramer test identified *XNT-S* to be distinct from the other three. Here, *XNT-S*, *Sticky Fingers*, and *XNT* were 84%, 12%, and 8% slower than *TNT-Touch* respectively. See Figure 8.

Technique had a significant main effect on position error ($F_{3,11} = 32.25, p < .0001$). A Tukey-Kramer test suggested *XNT-S* to be separate from all other techniques. Technique also had a significant main effect on orientation error ($F_{3,11} = 18.25$, $p < .0001$), and *TNT-Touch* was identified as distinct from the other techniques by Tukey-Kramer. Among the other three, *XNT* and *XNT-S* were distinct from each other, but not from *Sticky Fingers*.

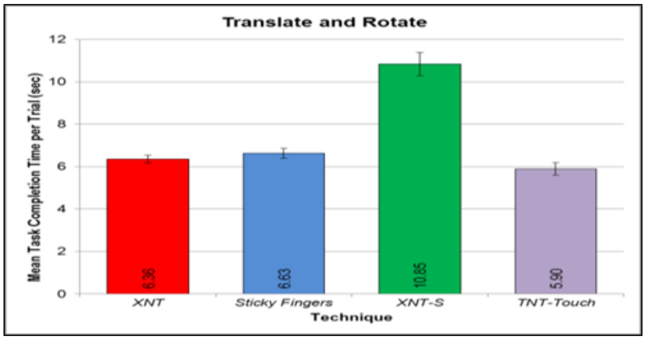

Figure 8: Mean task completion time by technique for the Translate and Rotate task. Error bars show ±1 *SE*.

A significant main effect also existed for technique on size error ($F_{2,11} = 16.02, p < .05$). Given that our implementation of *TNT-touch* could not resize objects, there were no sizing errors. *Sticky Fingers* and *XNT-S* were distinct from each other, but not from *XNT*. See Figure 9. No significant interaction existed between target orientation and technique ($F_{3,11} = 1.41, p > .05$).

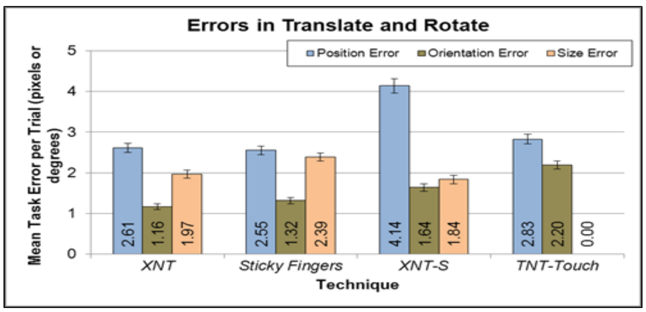

Figure 9: Mean task Error by technique for the Translate and Rotate task. Position and Size errors are reported in pixels and orientation errors in degrees. Error bars show ±1 *SE*.

6.3 Translate, Rotate, and Scale Task

A repeated-measure ANOVA on the data revealed a significant main effect of technique on task completion time ($F_{2,11} = 22.15, p < .0001$). The Tukey-Kramer test placed *XNT* in a group distinct from the others. Here, *XNT-S* and *Sticky Fingers* were 65% respectively 55% slower. See Figure 10.

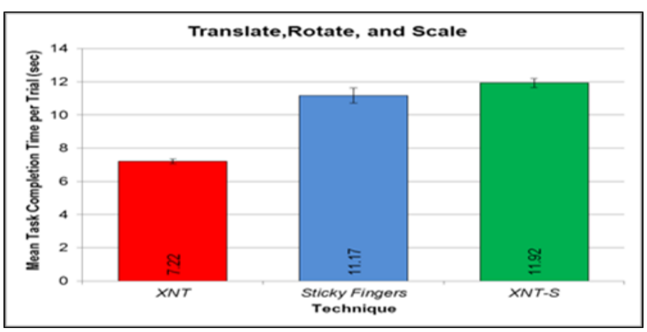

Figure 10: Mean task completion time by technique for the Translate, Rotate, and Scale task. Error bars show ±1 *SE*.

Technique had a significant main effect on position error ($F_{2,11} = 6.76, p < .01$) with *Sticky Fingers* and *XNT* distinct from each other, but not from *XNT-S*. A main significant effect was also found on orientation error ($F_{2,11} = 11.32, p < .0005$). *Sticky Fingers* is in a group distinct from the other two. The effect of technique on size error is also significant ($F_{2,11} = 13.77, p < .0005$). *XNT-S* was distinct from the other two techniques. See Figure 11.

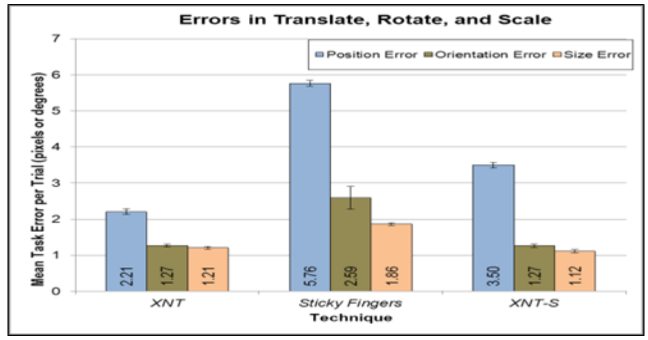

Figure 11: Mean task Error by technique for the Translate, Rotate, and Scale task. Position and Size errors are reported in pixels and orientation errors in degrees. Error bars show ±1 *SE*.

A significant interaction existed between technique and path ($F_{2,11} = 5.37$, $p < .0001$). In straight horizontal movements (i.e., AB and BA), higher performance was observed in left to right motion. In straight vertical movements (i.e., DC and CD), motion towards the person or downwards led to higher performance. In top-left to bottom-right diagonal movements (i.e., BC and CB), starting in the top-left was only slightly better than the reverse. In top-right to bottom-left diagonal movements (i.e., BD and DB), performance was higher in the bottom-left to top-right direction. DC had the worst average performance and AB the best.

A significant interaction also existed between technique and target orientation ($F_{2,11} = 2.83, p < .01$). *XNT* and *XNT-S* had a similar pattern in that they had higher performance when targets needed to be rotated by only 45 degrees compared to 135 degrees, with a slight bias for clockwise rotation. *Sticky Fingers* followed this pattern for clockwise rotations. However, it showed the opposite behaviour when targets were rotated counter-clockwise.

The interaction between technique and target size was significant ($F_{2,11} = 20.93, p < .0001$). Performance of the *XNT* and *XNT-S* techniques decreased approximately linearly as the target became smaller with *XNT* having a smaller slope. *Sticky Fingers* experienced the same relative performance hit when the target was half size; however, its performance dropped much stronger when manipulating the quarter size targets.

6.4 Questionnaire

In the questionnaire given to the participants, they were asked to select their preferred technique with regards to its performance, accuracy, fatigue, and overall usability.

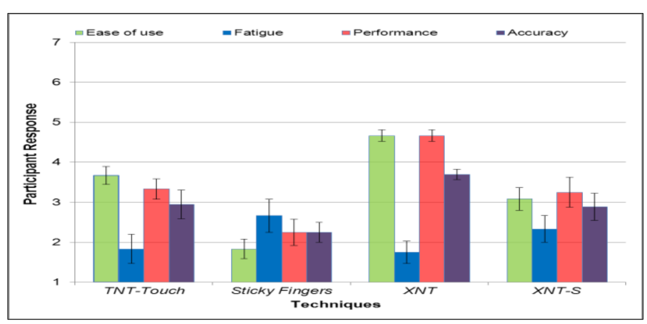

Figure 12: Results of the questionnaire for comfort and preference. A response of 7 is most-favourable, and 1 least favourable. Error bars = ±1 *SE*.

6.5 Main Results Calculated from First Touch Event

While the completion time results were computed starting from the appearance of the object to keep our study compatible with previous research [30], we also recorded completion time from the moment of first touch on the object. For the *Translate and Scale task*, a repeated-measure ANOVA revealed a significant main effect of technique on task completion time ($F_{2,11} = 4.33$, $p < .05$). *XNT*, *XNT-S*, and *Sticky Fingers* have mean completion times of 7.94, 8.76, and 15.14 seconds respectively. *XNT-S* and *Sticky Fingers* are 10% and 91% slower than *XNT* in this task. For the *Translate and Rotate task*, a repeated-measure ANOVA revealed a significant main effect of technique on task completion time ($F_{3,11} = 10.98$, $p < .0001$). *TNT-Touch, XNT, Sticky Fingers*, and *XNT-S* have mean completion times of 4.86, 5.30, 5.53, and 9.50 seconds. *XNT, Sticky Fingers*, and *XNT-S* are 9%, 14%, and 95% slower than *TNT-Touch*. For the *Translate, Rotate, and Scale task*, a repeated-measure ANOVA revealed a significant main effect of technique on task completion time ($F_{2,11} = 23.42$, $p < .0001$). *XNT, Sticky Fingers*, and *XNT-S* have mean completion times of 6.32, 10.14, and 10.91 seconds respectively. *Sticky Fingers* and *XNT-S* are 60% and 73% slower than *XNT*.

7 Discussion

As above, the discussions here are organized by task. Naturally, results for each technique are only valid for the tasks supported.

7.1 Translate and Scale Task

The results for this task suggest a positive correlation between performance and accuracy. *XNT* is the best performing technique and also the most accurate in terms of position and size. *Sticky Fingers* performs worse than the other two techniques and is the least accurate. During the experiment, participants were observed to struggle in trials that required matching the quarter size target. Ten participants voiced their discomfort and lack of control here. We believe this to be due to the requirement to have both fingers on the object to manipulate it in our implementation of *Sticky Fingers*. This reduces visibility and also makes it difficult to manipulate the object once it becomes small, as the fingers can not get too close without being recognized as a single touch.

The interaction between the technique and target size suggest that for targets of quarter size, *XNT-S* performs slightly better than *XNT* and is less prone to the performance drop caused by a small target size. *Sticky Fingers* is the least suitable method for small target sizes and preformed quite badly in this particular task.

XNT-S performs slightly worse than *XNT*, but is still reasonably accurate in terms of position and size. Since targets in this task were always the same orientation as the object, this technique had the potential for no orientation error. Yet, participants sometimes accidentally touched with their third finger or got confused on how many fingers to use to perform the desired action. This explains the existence of a non-zero orientation error in *XNT-S*.

7.2 Translate and Rotate Task

Here, we can observe a performance gap between integral and separated techniques. *XNT-S*, a separated technique, performs significantly worse than the others three and *XNT* in particular. It also has the worst position accuracy as a result of its rigidness in control and inability to perform simultaneous translation and rotation. *TNT-Touch* demonstrated the best performance at the expense of orientation accuracy, as small angular adjustments are more difficult to perform using the wrist than the fingers. Hence, we see this as evidence that *TNT-Touch* is a good technique for fast translation and rotation whenever orientation accuracy is not a primary concern, size manipulation is not required, and fingers can be held together while still being identifiable.

XNT is slightly faster than *Sticky Fingers* and has also the best performance in terms of size and orientation accuracy. *Sticky Fingers* is slightly better in terms of positioning accuracy, but none of these differences are statistically significant. Hence, we can recommend either of the two techniques for this task.

7.3 Translate, Rotate, and Scale Task

XNT yields again the best performance for this task. It also features the best position and orientation accuracy. *XNT-S* provides similar orientation accuracy and slightly better size accuracy than *XNT*, but at the expense of worse performance. *Sticky Fingers* is the least accurate technique and also the worst performer. In other words, for full-2D tasks, *XNT* provides the best balance of performance and accuracy.

XNT shows the most consistent performance across different paths and different orientations. It is also least prone to potential penalties due to smaller target sizes. *XNT-S* is the second-best technique in these regards. *Sticky Fingers* is most sensitive to path directions and rotations and cannot deal well with small targets.

7.4 Questionnaire

The participants' subjective responses agree with the objective performance results of the experiment in general. Participants perceived *XNT* to be the easiest technique to use and least fatiguing. We attribute this partly to the usage of a stable three-finger posture and partly to the selection-by-centroid feature of *XNT*. *XNT-S* shares these features, hence its' lower ratings may be attributed to a preference for integral over separated techniques.

Interestingly, *XNT-S* is perceived to perform better than *Sticky Fingers*, while the data shows this to be true only in the Translate and Scale task. It is possible that participants' opinion of ease of use and fatigue for *Sticky Fingers* may have affected their perception of how the technique performed. Also, since *TNT-Touch* was not tested in tasks that involved scaling, it is likely that participants became more familiar with the other techniques. This may have affected the ratings for this technique.

7.5 Overall discussion

Overall, *XNT* performed either best or equivalent to another technique across all tasks. Thus, we can recommend *XNT* as the best choice for general 2D object transformations. If no scaling is involved, *Sticky Fingers* or *TNT-Touch* are also acceptable. We hypothesized in advance that changing the calculation of completion time from the moment of object appearance to the moment of first touch would not affect the results of this study. Our analysis of the data confirms that, as none of the main results changed in any substantive manner. We see this as evidence that potentially different cognitive preparation times have likely no strong effect on the performance of the tested techniques.

While long-term practice may affect the performance and accuracy of these techniques, we chose not to investigate such learning in depth. Our focus was on the usability of these multi-touch techniques for the novice users. In general, techniques that are easy-to-learn, easy-to-use, yet efficient, contribute to quicker adoption of a technology, especially for novices. The result of this study can guide system designers to implement task-appropriate techniques that enhance the user experience on digital tabletops.

8 Conclusion

We presented two new three-finger object transformation techniques for multi-touch surfaces, *XNT* and *XNT-S*. We performed a study that evaluated these and several other multi-touch techniques in three object manipulation tasks on tabletop systems. The results reveal that the new *XNT* technique is an excellent choice for all kinds of manipulation tasks on multi-touch systems, especially if objects need to be scaled.

We focused our effort on similarity 2D object manipulations here, which is not sufficient for creating new content. In a 2D drawing application, features such as tearing the object or deforming parts of the image can promote creativity. Exploring multi-touch techniques for such transformations introduces new opportunities. More research is also needed to increase the precision of selecting a small target surrounded by multiple other densely-arranged objects.

9 REFERENCES

[1] Microsoft Surface. http://www.microsoft.com/surface.

[2] SmartTech http://www.smarttech.com

[3] DiamondTouch. http://www.circletwelve.com

[4] Perceptive Pixel. http://www.perceptivepixel.com

[5] Autodesk. http://usa.autodesk.com

[6] Baudel, T., and Beaudouin-Lafon, M. Charade: remote control of objects using free-hand gestures. CACM 36, 7 (1993), 28-35.

[7] Beaudouin-Lafon, M. Novel interaction techniques for overlapping windows. UIST 2001, 153-154.

[8] Benko, H., Wilson, A.D., and Baudisch, P. Precise selection techniques for multi-touch screens. CHI 2006, 1263-1272.

[9] Bier, E.A., Stone, M.C., Pier, K., Buxton, W., and DeRose, T.D. Toolglass and magic lenses: the see-through interface. SIGGRAPH 1993, 73-80.

[10] Epps, J., Lichman, S., and Wu, M. A study of hand shape use in tabletop gesture interaction. CHI EA 2006, 748-753.

[11] Fitzmaurice, G.W., Ishii, H., and Buxton, W.A.S. Bricks: laying the foundations for graspable user interfaces. CHI 1995, 442–449.

[12] Forlines, C., Shen, C., Vernier, F., and Wu, M. Under My Finger: Human Factors in Pushing and Rotating Documents Across the Table. INTERACT 2005, 994-997.

[13] Guiard, Y. Asymmetric division of labor in human skilled bimanual action: The kinematic chain as a model. Journal of Motor Behavior 19, 4 (1987), 486-517.

[14] Han, J.Y. Low-cost multi-touch sensing through frustrated total internal reflection. UIST 2005, 115-118.

[15] Hancock, M., Vernier, F., Wigdor, D., Carpendale, S., and Chia Shen. Rotation and translation mechanisms for tabletop interaction. TableTop 2006, 79-88 pp.

[16] Jacob, R.J.K., Sibert, L.E., McFarlane, D.C., and M. Preston Mullen, Jr. Integrality and separability of input devices. ACM Trans. Comput.-Hum. Interact. 1, 1 (1994), 3-26.

[17] Karat, J., McDonald, J.E., and Anderson, M. A comparison of menu selection techniques: touch panel, mouse and keyboard Int. J. Man-Mach. Stud. 25, 1 (1986), 73-88.

[18] Kin, K., Agrawala, M., and DeRose, T. Determining the benefits of direct-touch, bimanual, and multifinger input on a multitouch workstation. GI 2009, 119-124.

[19] Kruger, R., Carpendale, S., Scott, S.D., and Greenberg, S. Roles of orientation in tabletop collaboration: comprehension, coordination and communication. Computer Supported Cooperative Work, 13, 5-6 (2004), 501-537.

[20] Kruger, R., Carpendale, S., Scott, S.D., and Tang, A. Fluid integration of rotation and translation. CHI 2005, 601-610.

[21] Liu, J., Pinelle, D., Sallam, S., Subramanian, S., and Gutwin, C. TNT: improved rotation and translation on digital tables. GI 2006, 25-32.

[22] Mark Hancock, Thomas ten Cate, and Sheelagh Carpendale. Sticky tools: full 6DOF force-based interaction for multi-touch tables. ITS 2009, 133-140.

[23] Matsushita, N., Ayatsuka, Y., and Rekimoto, J. Dualtouch: a two-handed interface for pen-based PDAs. UIST 2000, 211-212.

[24] Micire, M., Desai, M., Courtemanche, A., Tsui, K.M., and Yanco, H.A. Analysis of natural gestures for controlling robot teams on multi-touch tabletop surfaces. ITS 2009, 41-48.

[25] Mitchell, G.D. Orientation on tabletop displays. M.Sc. Thesis, Simon Fraser University, Burnaby, BC, 2003.

[26] Morris, M.R., Wobbrock, J.O., and Wilson, A.D. Understanding users' preferences for surface gestures. GI 2010, 261-268.

[27] Moscovich, T. and Hughes, J.F. Indirect mappings of multi-touch input using one and two hands. CHI 2008, 1275-1284.

[28] Nacenta, M.A., Baudisch, P., Benko, H., and Wilson, A. Separability of spatial manipulations in multi-touch interfaces. GI 2009, 175-182.

[29] Oh, J. and Stuerzlinger, W. Laser pointers as collaborative pointing devices. Graphics Interface 2002, 141-149.

[30] Ostroff, D. and Shneiderman, B. Selection devices for user of an electronic encyclopedia: An empirical comparison of four possibilities. Inf. Process. Manage. 24, 6 (1988), 665-680.

[31] Reisman, J.L., Davidson, P.L., and Han, J.Y. A screen-space formulation for 2D and 3D direct manipulation. UIST 2009, 69-78.

[32] Rekimoto, J. and Saitoh, M. Augmented surfaces: a spatially continuous work space for hybrid computing environments. CHI 1999, 378-385.

[33] Ringel, M., Ryall, K., Shen, C., Forlines, C., and Vernier, F. Release, relocate, reorient, resize: fluid techniques for document sharing on multi-user interactive tables. CHI EA 2004, 1441-1444.

[34] Ryall, K., Forlines, C., Shen, C., and Morris, M.R. Exploring the effects of group size and table size on interactions with tabletop shared-display groupware. CSCW 2004, 284-293.

[35] Schmidt, D., Block, F., and Gellersen, H. A Comparison of Direct and Indirect Multi-touch Input for Large Surfaces. INTERACT 2009, 582-594.

[36] Shen, C., Vernier, F.D., Forlines, C., and Ringel, M. DiamondSpin: an extensible toolkit for around-the-table interaction. CHI 2004, 167-174.

[37] Streitz, N.A., Geißler, J., Holmer, T., et al. i-LAND: an interactive landscape for creativity and innovation. CHI 1999, 120-127.

[38] Stuerzlinger, W., Zaman, L., Pavlovych, A., and Oh, J. The design and realization of CoViD: a system for collaborative virtual 3D design. Virtual Real. 2006, 135-147.

[39] Tandler, P., Prante, T., Müller-Tomfelde, C., Streitz, N., and Steinmetz, R. Connectables: dynamic coupling of displays for the flexible creation of shared workspaces. UIST 2001, 11-20.

[40] Wang, Y., MacKenzie, C.L., Summers, V.A., and Booth, K.S. The structure of object transportation and orientation in human-computer interaction. CHI 1998, 312-319.

[41] Wigdor, D., Williams, S., Cronin, M., et al. Ripples: utilizing per-contact visualizations to improve user interaction with touch displays. UIST 2009, 3-12.

[42] Wu, M., Chia Shen, Ryall, K., Forlines, C., and Balakrishnan, R. Gesture registration, relaxation, and reuse for multi-point direct-touch surfaces. TableTop 2006, 8 pp.

[43] Wu, M. and Balakrishnan, R. Multi-finger and whole hand gestural interaction techniques for multi-user tabletop displays. UIST 2003, 193-202.