Understanding Multi-touch Manipulation
for Surface Computing

Chris North[1], Tim Dwyer[2], Bongshin Lee[2], Danyel Fisher[2], Petra Isenberg[3],
George Robertson[2], Kori Inkpen Quinn[2]

[1] Virginia Tech, Blacksburg, VA, USA {north@vt.edu}
[2] Microsoft Research, Redmond, WA, USA
{t-tdwyer,danyelf,bongshin,ggr,kori@microsoft.com}
[3] University of Calgary, Alberta, Canada {petra.isenberg@ucalgary.ca}

Abstract. Two-handed, multi-touch surface computing provides a scope for interactions that are closer analogues to physical interactions than classical windowed interfaces. The design of natural and intuitive gestures is a difficult problem as we do not know how users will approach a new multi-touch interface and which gestures they will attempt to use. In this paper we study whether familiarity with other environments influences how users approach interaction with a multi-touch surface computer as well as how efficiently those users complete a simple task. Inspired by the need for object manipulation in information visualization applications, we asked users to carry out an object sorting task on a physical table, on a tabletop display, and on a desktop computer with a mouse. To compare users' gestures we produced a vocabulary of manipulation techniques that users apply in the physical world and we compare this vocabulary to the set of gestures that users attempted on the surface without training. We find that users who start with the physical model finish the task faster when they move over to using the surface than users who start with the mouse.

Keywords: Surface, Multi-touch, Gestures, Tabletop

1 Introduction

The rapidly-developing world of multi-touch tabletop and surface computing is opening up new possibilities for interaction paradigms. Designers are inventing new ways of interacting with technology and users are influenced by their previous experience with technology.

Tabletop gestures are an important focal point in understanding these new designs. Windowing environments have taught users to experience computers with one hand, focusing on a single point. What happens when those constraints are relaxed, as in multi-touch systems? Does it make sense to allow—or expect—users to interact with multiple objects at once? Should we design for users having two hands available for their interactions? Both the mouse-oriented desktop and the physical world have constraints that limit the ways in which users can interact with multiple objects and users come to the tabletop very accustomed to both of these.

There is no shortage of applications where users might need to manipulate many objects at once. From creating diagrams to managing files within a desktop metaphor, users need to select multiple items in order to move them about. A number of projects in the visual analytics [11] and design spaces [6] have attempted to take advantage of spatial memory by simulating sticky notes—a mixed blessing when rearranging the notes is expensive and difficult. As it becomes simpler to move objects and the mapping between gesture and motion becomes more direct, spatial memory can become a powerful tool.

We would like to understand what tools for managing and manipulating objects the tabletop medium affords and how users respond to it. Particularly, we would like to understand the techniques that users adopt to manipulate multiple small objects. What techniques do they use in the real world and how do those carry over to the tabletop context? Do they focus on a single object—as they do in the real world—or look at groups? Do they use one hand or two? How dexterous are users in manipulating multiple objects at once with individual fingers?

The problems of manipulating multiple objects deftly are particularly acute within the area of visual analytics [13], where analysts need to sort, filter, cluster, organize and synthesize many information objects in a visualization. Example systems include In-Spire [16], Jigsaw [12], Occulus nSpace [10], or Analyst's Notebook [4], i.e. systems where analysts use virtual space to organize iconic representations of documents into larger spatial representations for sensemaking or presenting results to others. In these tasks, it is important to be able to efficiently manipulate the objects and it is often helpful to manipulate groups of objects. Our general hypothesis is that multi-touch interaction can offer rich affordances for manipulating a large number of objects, especially groups of objects.

A partial answer to these questions comes from recent work by Wobbrock et al. [17]. Users in that study were asked to develop a vocabulary of gestures; the investigators found that most (but not all) of the gestures that users invented were one-handed. However, their analysis emphasized manipulating single objects: they did not look at how users would handle gestures that affect groups of items.

In this paper we explore how users interact with large numbers of small objects. We discuss an experiment in which we asked users to transition from both a mouse and a physical condition to an interactive surface, as well as the reverse. We present a taxonomy of user gestures showing which ones were broadly used and which were more narrowly attempted. We also present timing results showing that two-handed tabletop operations can be faster than mouse actions, although not as fast as physical actions. Our research adds a dimension to Wobbrock et al.'s conclusions showing that two-handed interaction forms a vital part of surface gesture design.

2 Background

Typical interactions on groups of items in mouse-based systems first require multi-object selection and then a subsequent menu selection to specify an action on the selected objects. Common techniques for multi-object selection include drawing a selection rectangle, drawing a lasso, or holding modifier keys while clicking on several objects. In gestural interfaces this two-step process can be integrated into one motion. Yet, the design of appropriate gestures is a difficult task: the designer must develop gestures that can be both reliably detected by a computer and easily learned by people [5].

Similar to the mouse, pen-based interfaces only offer one point of input on screen but research on pen gestures is relatively advanced compared to multi-touch gestures. Pen-based gestures for multiple object interaction have, for example, been described by Hinckley et al. [3]. Through a combination of lasso selection and marking-menu-based command activation, multiple targets can be selected and a subsequent action can be issued. A similar example with lasso selection and subsequent gesture (e.g., a pigtail for deletion) were proposed for Tivoli, an electronic whiteboard environment [9].

For multi-touch technology, a few gesture sets have been developed which include specific examples of the types of multi-object gestures we are interested in. For example, Wu et al. [18] describe a Pile-n-Browse gesture. By placing two hands on the surface, the objects between both hands are selected and can be piled by scooping both hands in or browsed through by moving the hands apart. This gesture received a mixed response in an evaluation. Tse et al. [14] explore further multi-touch and multi-modal group selection techniques. To select and interact with multiple digital sticky notes, users can choose between hand-bracketing (similar to [18]), single-finger mouse-like lasso-selection, or a speech-and gesture command such as "search for similar items." Groups can then be further acted upon through speech and gestures. For example, groups of notes can be moved around by using a five-fingered grabbing gesture and rearranged through a verbal command. Using a different approach, Wilson et al. [15] explore a physical based interaction model for multi-touch devices. Here, multiple objects can be selected by placing multiple fingers on objects or by "pushing" with full hand shapes or physical objects against virtual ones to form piles.

Many of the above multi-selection gestures are extremely similar to the typical mouse-based techniques (with the notable exception of [15]). Wobbrock et al. [17] present a series of desired effects, and invite users to "act out" corresponding gestures in order to define a vocabulary. Participants described two main selection gestures—tap and lasso—for both single and group selection. This research also showed a strong influence of mouse-based paradigms in the gestures participants chose to perform. Similarly, our goal was to first find out which gestures would be natural choices for information categorization and whether a deviation from the traditional techniques of lasso or selection rectangles would be a worthwhile approach.

Previous studies have examined the motor and cognitive effects of touch screens and mouse pointers, and the advantages of two-handed interaction over one-handed techniques, primarily for specific target selection tasks (e.g., [1,7]). Our goal is to take a more holistic view of multi-touch interaction in a more open-ended setting of manipulating and grouping many objects.

3   Baseline Multi-touch Surface Interaction

Our goal is to study tasks in which users manipulate large numbers of small objects on screen. For our study, we abstracted such analytic interactions with a task involving sorting colored circles in a simple bounded 2D space.

Our study tasks, described below, involved selecting and moving colored circles on a canvas. We were particularly interested in multi-touch support for single and group selection of such objects. To provide a study platform for comparison with standard mouse-based desktop and physical objects conditions, we had to make some interaction design decisions for our baseline multi-touch system. Our design incorporates several assumptions about supporting object manipulation for surface computing:

• One or two fingers touching the surface should select individual objects.
• A full hand, or three or more fingers touching the surface, should select groups of objects.
• Contacts far apart probably indicate separate selections (or accidental contact) instead of a very large group. Unintentionally selecting a large group is more detrimental than selecting small groups.
• Multiple contacts that are near each other but initiated at different times are probably intended to be separate selections. Synchronous action might indicate coordinated intention.

The system (Fig. 1) is implemented on the Microsoft Surface [8], a rear-projection multi-touch tabletop display. The Surface Software Development Kit provides basic support for hit testing of users' contact points on the display. It also provides coordinates and an ellipsoidal approximation of the shape of the contact, as well as contact touch, move, and release events.

Our testing implementation supports selecting and dragging small colored circles both individually and in groups. The interaction design was intentionally kept simple to support our formative study goals. Contacts from fingers and palms select all the circles within their area. As feedback of a successful selection, the circles are highlighted by changing the color of their perimeters, and can be dragged to a new position. From there, they can be released and de-selected. A (small) fingertip contact selects only the topmost circle under the contact, enabling users to separate overlapping circles. Large contacts such as palms select all circles under the contact. Using multiple fingers and hands, users can manipulate multiple circles by such direct selection and move them independently. Such direct selection techniques are fairly standard on multi-touch interfaces.
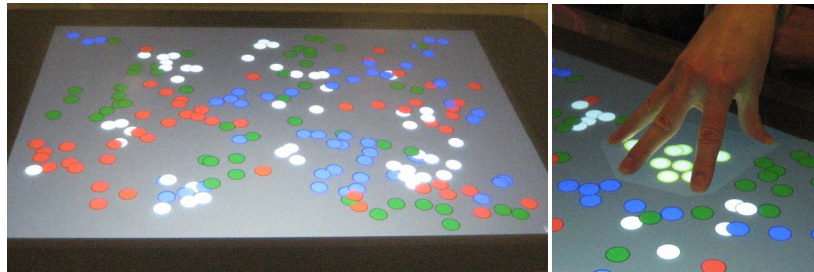

Fig. 1. Left: The system at the start of Task 1. **Right**: One-handed hull selection technique.

We also provide an analogue to the usual mouse-based rectangular marquee selection of groups of objects. However, a simple rectangular marquee selection does not make effective use of the multi-touch capability. Instead, the users can multi-select by defining a convex hull with three or more fingers. If three or more contacts occur within 200ms and within a distance of 6 inches from each other (approximately a hand-span), then a convex-hull is drawn around these contacts and a "group" selection is made of any circles inside this hull (Fig. 1, right). The background area inside the hull is also colored light grey to give the user visual feedback. These hulls, and the circles within them, can then be manipulated with affine transformations based on the users' drag motions. For example, users can spread out or condense a group by moving their fingers or hands together or apart. While the group selection is active, users can grab it with additional fingers to perform the transformations as they desire. The group selection is released when all contacts on the group are released.

4   Study Design

The goal of this study is to discover how users manipulate many small objects, in three different interaction paradigms: physical, multi-touch, and mouse interaction. To support our formative design goals, we took a qualitative exploratory approach with quantitative evaluation for comparisons.

4.1   Participants

We recruited 32 participants (25 males and 7 females) and 2 pilot testers via email from our institution. We screened participants for color blindness. They were mainly researchers and software developers who were frequent computer users. The average age of participants was 34, ranging from 21 to 61. None of the participants had significant experience using the Surface. Participants had either never used the Surface before or had tried it a few times at demonstrations. Participants each received a $US10 lunch coupon for their participation. To increase motivation, additional $10 lunch coupons were given to the participants with the fastest completion time for each interface condition in the timed task.

4.2   Conditions and Groups

We compared three interface conditions: Surface, Physical and Mouse. For both the Surface and Physical conditions, we used a Microsoft Surface system measuring 24" × 18". For the *Surface* condition (Fig. 1, left), we ran the multi-touch implementation described in Section 3 with 1024 × 768 resolution. For the *Physical* condition (Fig. 2, Left), we put 2.2cm diameter circular plastic game chips on top of the Microsoft Surface tabletop with same grey background (for consistency with the Surface condition). The circles in the Surface condition were the same apparent size as the game chips in the Physical condition.
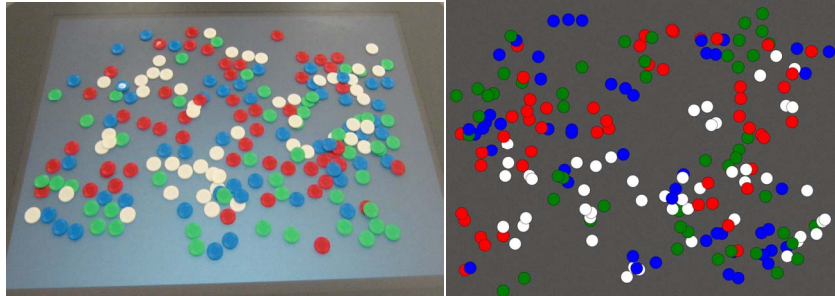
Fig. 2. (Left) Physical condition and **(Right)** Mouse condition.

For the *Mouse* condition (Fig. 2, right), we ran a C# desktop application on a 24" screen. This application supported basic mouse-based multi-selection techniques: marquee selection by drawing a rectangle as well as control- and shift-clicking nodes. Circles were sized so that their radii as a proportion of display dimensions were the same on both the desktop and surface.

Since our goal is to compare the Surface condition against the other two conditions, each participant used only two conditions: Surface and one of the others. Users were randomly divided into one of four groups: Physical then Surface (PS), Surface then Physical (SP), Mouse then Surface (MS), Surface then Mouse (SM). This resulted in participants' data for 32 Surface, 16 Physical and 16 Mouse.

4.3   Tasks

Participants performed four tasks, each task requiring spatially organizing a large number of small objects. The first and second tasks were intended to model how analysts might spatially cluster documents based on topics, and manage space as they work on a set of documents, and were designed to capture longer-term interaction strategies. The tasks required a significant amount of interaction by the participants and gave them a chance to explore the interface.

All participants worked on the four tasks in the same order, and were not initially trained on the surface or our application. Participants were presented with a table of 200 small circles, with 50 of each color: red, green, blue, and white. Fig. 1 illustrates the 200 circles on a Surface at the start of the first task, positioned randomly in small clusters.

With the exception of Task 3, which was timed, we encouraged participants to think aloud while performing the tasks so that we could learn their intentions and strategies.

Task 1: Clustering task. This task was designed to elicit users' intuitive sense of how to use gestures on the surface. The task was to organize the blue and white circles into two separate clusters that could be clearly divided from all others. Participants were told that the task would be complete when they could draw a line around the cluster without enclosing any circles of a different color. Fig. 3. shows one possible end condition of Task 1.
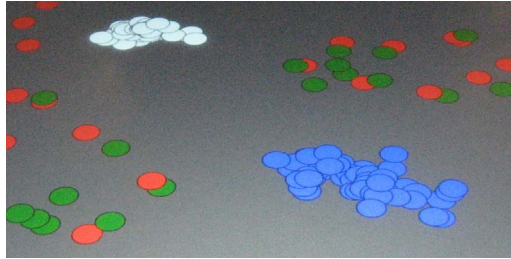
Fig. 3. Example end condition of Task 1.

Task 2: Spreading Task. Participants spread out the blue cluster such that no blue circles overlap, moving other circles to make room as needed. Participants start this task with the end result of their Task 1.

Task 3: Timed Clustering Task. This task was designed to evaluate user performance time for comparison between interface conditions and to examine the strategies which users adopt over time. Task 3 repeated Task 1, but participants were asked to complete the task as quickly as possible. They were not asked to 'think aloud' and a prize was offered for the fastest time.

Task 4: Graph Layout Task. Inspired by the recent study of van Ham and Rogowitz [2], we asked participants to lay out a social network graph consisting of 50 nodes and about 75 links. In the Physical condition, participants did not attempt this task. Due to the broader scope and complexity of this task, the analysis of the results of Task 4 will be reported elsewhere.

4.4   Procedure

Each participant was given an initial questionnaire to collect their demographics and prior experience with the Microsoft Surface system. Participants completed Tasks 1 and 2 without training, in order to observe the gestures they naturally attempted. Participants in the Surface and Mouse condition were given a brief tutorial about the available interaction features after Task 2. At the end of each condition participants answered a questionnaire about their experience. They then repeated the same procedure with the second interface condition. At the end of the session participants answered a final questionnaire comparing the systems. Each participant session lasted at most an hour.

We recorded video of the participants to capture their hand movements, their verbal comments and the display screen. The software also recorded all events and user operations for both the Surface and Mouse conditions.

5   RESULTS

We divide our results into an analysis of the set of gestures that users attempted for Tasks 1 and 2, timing results from Task 3, and user comments from the after-survey.

5.1   Gestures

The video data for Task 1 and 2 (clustering and spreading) were analyzed for the full set of operations users attempted in both the Physical and Surface conditions. We first used the video data to develop a complete list of all gestures, both successful and unsuccessful. For example, if a participant attempted to draw a loop on the surface, we coded that as an unsuccessful attempt to simulate a mouse "lasso" gesture. The gestures were aggregated into categories of closely-related operations. Once the gestures were identified, the videos were analyzed a second time to determine which gestures each user attempted.

Table 1 provides a listing of all classes of gestures that participants performed during the study; six of them are illustrated in Fig. 4. These gestures are divided into several categories: single-hand operations that affect single or groups of objects, two-handed gestures that affect multiple groups of objects, and two-handed gestures that affect single groups. Last, we list gestures that apply only to one medium: just surface, and just physical.

In order to understand how gestures varied by condition, we classed gestures by which participants attempted them. Table 2 lists all of the gestures that were feasible in both the Physical and Surface conditions. This table also lists the percentage of participants who utilized each gesture at least once during the session. This data is aggregated by the Physical and Surface conditions, followed by a further classification by which condition was performed first (Physical, Mouse, Surface). Table 3 lists additional gestures that were only feasible for the Surface condition, while Table 4 lists gestures that were only used in the Physical condition.

(a) One-hand shove.

(b) Drag two objects with pointer fingers.

(c) Two hands grab groups.

(d) Add/remove from selection.

(e) Two-hand transport.
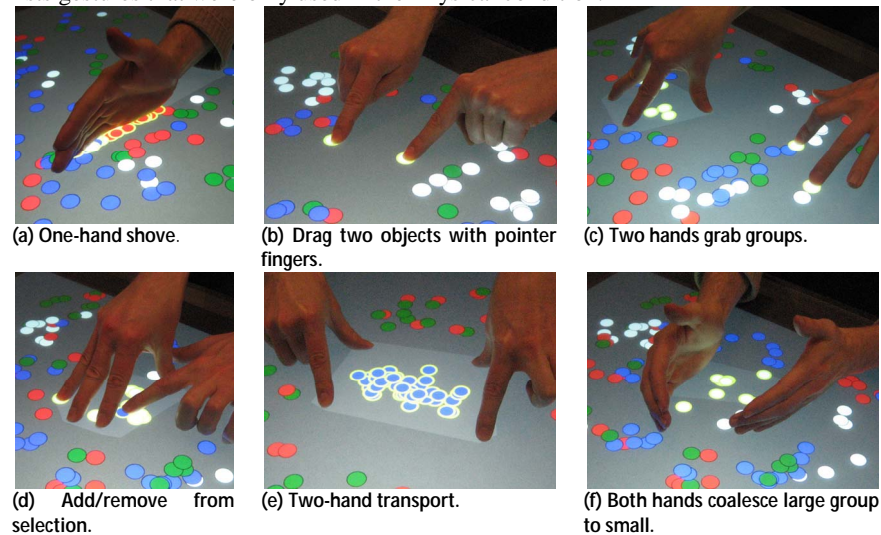
(f) Both hands coalesce large group to small.

Fig. 4. Six selected one- and two-handed gestures attempted by participants during the study (see Table 1 for the full list).

Table 1. Descriptions of gestures.

	Individual Items	Groups
One hand	**Drag single object**. Drag a single item across the tabletop with a fingertip. **Drag objects with individual fingers.** Using separate fingers from one hand, drag individual items across the table **Toss single object**. Use momentum to keep an object moving across the tabletop.	**Splayed hand pushes pieces**. (Fig. 1) An open hand pushes pieces. Could define a hull. **Hand and palm**. A single hand is pressed flat against the table to move items underneath it. **One hand shove** (Fig. 4a). Moves many points as a group. **Pinch a pile.** Several fingers "pinch" a group of pieces together. In the Surface condition, this would define a (very small) hull.

	Coordinated, >1 Group	Coordinated, 1 Group
Two hands	**Drag two objects with pointer fingers** (Fig. 4b). Does not entail any grouping operations. **Two hands grab points in sync**. Each hand has multiple fingers pulling items under fingers. **Rhythmic use of both hands.** "Hand-over-hand" and synchronized motion, repeated several or many times. **Two hands grab groups** (Fig 4c). Hands operate separately to drag groups or individual points.	**Both hands coalesce large group to small** (Fig 4f). **Two-hand transport** (Fig 4e). Use two hands to grab a group and drag across the region. **Add/remove from selection** (Fig 4d). Use one hand to pull an object out of a group held by the other.

	Surface only	Physical Gestures
By condition	**One hand expand/contract**. Use a single hand with a convex hull to grow or shrink a hull. **Two-hand hull tidy**. Use fingers from two hands with a convex hull to shrink the hull to make more space. **Two-hand hull expand/contract**. Use fingers from two hands with a convex hull to manipulate the hull. **Expand hull to cover desired nodes**. Define a hull first, then expand it to cover more nodes. Does not work on our Surface implementation. **Treat finger like a mouse**. Includes drawing a lasso or marquee with one or two hands, different fingers of the hand for "right" click, or holding down one hand to "shift-click" with the other. **Push hard to multi-select**. Press a finger *harder* into the table to hope to grow a selection or select more items in the near vicinity.	**Lift Up**. Pick up chips in the hand, carry them across the surface, and deposit them on the other side. **Go outside the lines**. Move, stack, or slide chips on the margin of the table, outside the screen area. **Slide around objects**. When sliding circles, choose paths across the space that avoid other circles. **"Texture"-based gestures.** Slide chips under palms and fingers and shuffle them, using the feel of the chip in the hand. **Toss items from one hand to other**. Take advantage of momentum to slide chips from one hand to the other. **Drag a handful, dropping some on the way**. Intentionally let some chips fall out of the hand, holding others, to either spread out a pile or sort them into different groups.

Across all participants the most popular gestures were those that entailed using fingertips to move circles across the table—all participants moved at least some items around that way. While all participants realized they could move physical objects with two hands, six of them never thought to try that with the Surface (three that started in Surface condition; three from group MS). Closer examination of the gesture data revealed that participants who started with the physical condition were much more

likely (88%) to try multiple fingers with both hands than users who started with the mouse (56%) or the surface (50%).

When participants worked with two hands on the surface they almost always used them on separate groups: only 30% of participants performed operations that used both hands at once to affect a single group. However, both hands were often used to move groups separately.

We observed several habits from the other conditions that crept into the Surface interactions. For example, 56% of users tried to use their fingers as a mouse, experimenting using a different finger on the same hand for a "multi-select" or trying to draw marquees or lassos. Half of the users who started with the mouse continued to try mouse actions on the surface, while 25% of users who started with the physical condition tried mouse actions. More results are summarized in Table 3.

Table 2. Gestures that apply to both the physical and surface conditions. Values indicate the percentage of subjects who used the gesture at least once. Values over 50% are boldfaced.

| | Physical | Surface | Surface (by 1st Condition) | | |
| | | | After Mouse | After Physical | Surface 1st |
	(n=16)	(n=32)	(n=8)	(n=8)	(n=16)
1 Hand, Individual Items					
Drag single object	**75%**	**94%**	**100%**	**75%**	**100%**
Drag objects with indiv fingers	**81%**	**69%**	50%	50%	**88%**
Toss single object	38%	19%	0%	13%	31%
1 Hand, Groups					
Splayed hand pushes pieces (Fig 1)	50%	28%	25%	25%	31%
One hand shove (Fig 3a)	**75%**	47%	38%	38%	**56%**
Hand and palm	31%	41%	25%	25%	**56%**
Pinch a pile	6%	38%	13%	25%	**56%**
2 Hands, Coordinated, > 1 Group					
Drag 2 objects with pointer fingers (3b)	**63%**	**63%**	50%	**88%**	**56%**
Two hands grab points in sync	**88%**	50%	38%	**88%**	38%
Rhythmic use of both hands	**56%**	41%	50%	**63%**	25%
Both hands grab groups (3c)	**81%**	**34%**	38%	50%	25%
2 Hands, Coordinated, 1 Group					
Both hands coalesce large group to small (3f)	**75%**	9%	13%	13%	6%
Two-hand transport (3e)	**69%**	41%	38%	63%	31%
Add/remove from selection (3d)	25%	19%	0%	13%	31%

Table 3. Gestures that apply only to Surface condition.

| | Surface (by 1st Condition) | | |
| | Mouse 1st | Physical 1st | Surface 1st |
	(n=8)	(n=8)	(n=16)
Hull Resizing			
One Hand Hull Expand/Contract	13%	13%	25%
Two hand hull Tidy	0%	25%	6%
Two-hand hull Expand/Contract	25%	63%	56%
Expand hull to cover desired nodes (doesn't work)	13%	25%	6%
Other (failures)			
Treat finger like a mouse	50%	25%	38%
Push hard to multi-select	25%	13%	31%

We wanted to understand what additional physical operations might be applied to a digital representation. In Table 4, we list operations that users performed in the physical condition that do not have a direct digital analogue. For example, 75% of all participants in the physical condition lifted the chips off the table; and 69% also pushed chips outside of the bounds of the table. Some of these gestures were attempted in the surface condition, but participants quickly realized that they were not supported on the surface. The one exception to this was a gesture to *slide objects around other objects* when moving them, which *was* possible in the surface condition although it was unnecessary since selected circles could be dragged *through* unselected circles.

Table 4. Gestures that refer to the Physical condition.

	Physical (n=16)	Surface (n=32)
Physical Gestures		
Lift Up	**75%**	3%
Go outside the lines	**69%**	0%
Slide *around* objects	**88%**	34%
"Texture"-based gestures (e.g. flattening a pile)	44%	3%
Toss items from one hand to other	38%	0%
Drag a handful, dropping some on the way	25%	6%

5.2   Timing Results for Task 3

In addition to articulating the set of possible operations, we also wanted to understand which ways of moving multiple objects were most efficient. Do participants do better with the two-handed grouping operations of the surface, or the familiar mouse? We analyzed the task time data with a 2 (Condition) × 2 (Group) mixed ANOVA. Table 5 shows mean completion times with standard deviations for Task 3.

Table 5. Mean completion times (with std deviations) for Task 3 in seconds.

Condition \ Order	MS (n=8)	PS (n=8)	SM (n=8)	SP (n=8)
Physical	-	71.0 (14.5)	-	107.6 (13.8)
Mouse	123.9 (30.9)	-	144.5 (32.5)	-
Surface	116.7 (21.8)	94.9 (30.3)	118.7 (31.2)	146.4 (37.5)

Surface is faster than Mouse. For the 16 participants who completed the Surface and Mouse conditions, we ran a 2 x 2 mixed ANOVA with *condition* {Surface, Mouse} as the within subjects variable and *order of conditions* as the between subjects variable. A significant main effect of condition was found ($F_{1,14}$=6.10, p=.027) with the surface condition being significantly faster (116 sec) than the mouse condition (134 sec). No significant effect of order was found ($F_{1,14}$=.928, p=.352) and there was no interaction effect between condition and order ($F_{1,14}$=1.38, p=.260).

Physical is faster than Surface, and trains users to be faster. For the 16 participants who completed the Surface and Physical conditions we again ran a 2 x 2 mixed ANOVA with *condition* {Surface, Physical} as the within subjects variable and *order of conditions* as the between subjects variable. A significant main effect of condition was found ($F_{1,14}$=11.96, p=.004) with the physical condition being significantly faster

(89 sec) than the surface condition (120 sec). In addition, a significant effect of condition order was found ($F_{1,14}$=11.482, p<.001) where participants who started with the physical condition were significantly faster than the participants who started with the surface condition. No significant interaction effect was found between condition and order ($F_{1,14}$=0.655, p=.432).

Impact of First Condition. We hypothesized that users' performance on the surface would be impacted by whether they started with the Mouse condition first, or the Physical condition first. Two participants data were classified as outliers (> Average + 1.5 * SD). An independent samples t-test revealed that participants who performed the physical condition first were significantly faster on the surface, than participants who performed the mouse condition first (t_{12}=2.38, p=.035).

Number of Group Operations. In attempting to understand the time difference reported in the previous section, we found that the physical-surface (PS) group used more group operations than the mouse-surface (MS) group: an average of 33 group operations across participants in the PS group against 26 for the MS group. However, this difference was not significant enough to reject a null-hypothesis (t_{14}=0.904, p=.381). Of course multi-touch interaction on the surface affords a number of other types of interaction that may increase efficiency in such a clustering task, e.g., simultaneous selection with multiple fingers or independent "hand-over-hand" gestures.

5.3 User Comments

We asked participants to rate the difficulty of the clustering task on a 7 point Likert scale (1=Very difficult, 7=Very easy). We ran a 2 x 2 mixed ANOVA with *condition* as the within subjects variable and *order of conditions* as the between subjects variable. We found a significant main effect of condition ($F_{1,14}$=5.8, p=.03) with the surface condition being significantly easier (5.5) than the mouse condition (4.9). No significant effect was found between Physical and Surface.

Participants seemed to appreciate the manipulation possibilities of the Surface: - when we asked which condition they preferred to perform the clustering task, 14 participants (88%) prefer Surface to Mouse. However, only 7 (44%) prefer Surface to Physical. Interestingly, the performance advantage of the Surface over the Mouse was greater than some participants thought. When we asked which condition they felt faster, only 9 participants (56%) felt Surface was faster than Mouse even though 12 (75%) actually did perform faster with Surface. However, 4 participants (25%) felt Surface was faster than Physical even though 3 (19%) were actually faster with Surface.

In verbal comments from participants who used both Physical and Surface, the most commonly cited advantage of Physical was the tactile feedback, i.e. selection feedback by feel rather than visual highlights. Whereas, the most cited advantage of the Surface was the ability to drag selected circles through any intervening circles instead of needing to make a path around them. For the participants who used both Mouse and Surface, the most cited advantage of the Mouse was multi-selecting many dispersed circles by control-clicking, while the most cited advantage of Surface was the ability to use two hands for parallel action.

6   Discussion and Conclusions

Tabletop multi-touch interfaces such as the Microsoft Surface present new opportunities and challenges for designers. Surface interaction may be more like manipulating objects in the real world than indirectly through a mouse interface, but it still has important differences from the real world, with its own advantages and disadvantages.

We observed that participants use a variety of two handed coordination. Some participants used two hands simultaneously, some used two hands in sync (hand over hand), some used coordinated hand-offs, and others used some combination of these. As a result, defining a group-by gesture requires some care because participants have different expectations about how grouping may be achieved when they first approach the Surface. In our particular implementation, participants sometimes had difficulty working with two hands independently and close together when our heuristic would make a group selection. We caution future designers of tabletop interfaces to consider this complexity in finding a good balance between physical metaphors and supporting gestures to invoke automation.

Multi-touch grouping turned out to be very useful. Many participants manipulated groups, and seemed to do so without thinking about it explicitly. Possibly the most valuable and common type of group manipulations were ephemeral operations such as the small open-handed grab and move. Massive group operations, such as moving large piles, also helped participants efficiently perform the clustering task. While our current implementation of group-select worked reasonably well as a baseline, we observed some difficulty with our hull system. We believe a better implementation of group select and increased user familiarity with multi-touch tabletop interfaces may bring user efficiency closer to what we observed in the Physical condition.

We have introduced a particular task that may be a useful benchmark for testing the efficiency and ergonomics of a particular type of basic tabletop interaction, but there is a great deal of scope for further studies. As was briefly mentioned in this paper, our study included a more challenging and creative task involving a layout of a network diagram. We intend to follow-up on this first exploration with an evaluation of a user-guided automatic layout interface that attempts to exploit the unique multi-touch capability of tabletop systems.

References

1. Forlines, C., Wigdor, D., Shen, C., Balakrishnan, R.: Direct-Touch vs. Mouse Input for Tabletop Displays. Proc. CHI '07, pp. 647-656. ACM Press, New York (2007)
2. Ham, F. van., Rogowitz, B.: Perceptual Organization in User-Generated Graph Layouts. IEEE Trans. Visualization and Computer Graphics (InfoVis '08), 14(6), pp. 1333-1339. IEEE Press, Los Alamitos, CA, USA (2008)

3. Hinckley, K., Baudisch, P., Ramos, G., Guimbretière, F.: Design and Analysis of Delimiters for Selection-action Pen Gesture Phrases in Scriboli. Proc. CHI '05, pp. 451-460. ACM Press, New York, NY, USA (2005)
4. i2 – Analyst's Notebook. http://www.i2inc.com (accessed 29 January 2009)
5. Jong, Jr. A. C., Landay J. A., Rowe L. A: Implications for a Gesture Design Tool. Proc. CHI '99, pp. 40-47. ACM Press, New York, NY, USA (1999)
6. Klemmer, S.R., Newman, M.W., Farrell, R. Bilezikjian, M., Landay., J.A.: The Designers' Outpost: A Tangible Interface for Collaborative Web Site Design. Proc. UIST '01, pp. 1-10. ACM Press, New York, NY, USA (2001)
7. Leganchuk, A., Zhai, S., Buxton, W.: Manual and Cognitive Benefits of Two-Handed Input: An Experimental Study. ACM Transaction on Computer-Human Interaction, 5(4), pp. 326-359. ACM Press, New York, NY, USA (1998)
8. Microsoft Surface, http://www.microsoft.com/surface
9. Pedersen, E.R., McCall, K., Moran, T. and Halasz, F. Tivoli: An Electronic Whiteboard for Informal Workgroup Meetings. Proc. CHI '93, pp. 391-398. ACM Press, New York, NY, USA (1993)
10. Proulx, P., Chien, L., Harper, R., Schroh, D., Kapler, T., Jonker, D., Wright, W.: nSpace and GeoTime: A VAST 2006 Case Study, IEEE Computer Graphics and Applications, 27 (5), pp. 46--56. IEEE Press, Los Alamitos, CA, USA (2007)
11. Robinson, Anthony C. Collaborative Synthesis of Visual Analytic Results. Proc. VAST '08, pp. 61-74. IEEE Press, Los Alamitos, CA, USA (2008)
12. Stasko, J., Görg, C., Liu, Z.: Jigsaw: supporting investigative analysis through interactive visualization. Information Visualization, 7 (2), pp. 118–132, Palgrave Macmillan, Houndmills, Basingstoke, Hampshire, England (2008)
13. Thomas, J.J., Cook, K.A.: Illuminating the Path. IEEE Press, Los Alamitos, CA, USA (2005)
14. Tse, E., Greenberg, S., Shen, C., Forlines, C., Kodama, R.: Exploring true multi-user multimodal interaction over a digital table. Proc. DIS '08, pp. 109-118. ACM Press, New York, NY, USA (2008)
15. Wilson, A.D., Izadi, S., Hilliges, O., Garcia-Mendoza, A., Kirk, D: Bringing physics to the surface. Proc. UIST '08, pp. 67--76. ACM Press, New York, NY, USA (2008)
16. Wise, J.A., Thomas, J.J., Pennock, K., Lantrip, D., Pottier, M., Schur, A., Crow, V.: Visualizing the non-visual: spatial analysis and interaction with information from text documents. Proc. InfoVis '95, pp. 51-58. IEEE Press, Los Alamitos, CA, USA USA (1995)
17. Wobbrock, J., Morris M.R., Wilson, D.A.: User-Defined Gestures for Surface Computing. Proc. CHI '09, to appear. ACM Press, New York, NY, USA (2009)
18. Wu, M., Shen, C., Ryall, K., Forlines, C., Balakrishnan, R.: Gesture Registration, Relaxation, and Reuse for Multi-Point Direct-Touch Surfaces. Proc. TableTop '06, pp.185-192. IEEE Press, Los Alamitos, CA, USA (2006)