Author manuscript, published in International Journal of Human-Computer Studies, 78(0), pp. 68-80, Jun 2015.
DOI: 10.1016/j.ijhcs.2015.02.011

Exploring and Modeling Unimanual Object Manipulation on Multi-Touch Displays

Jian Zhao*, R. William Soukoreff, Ravin Balakrishnan

Department of Computer Science, Univeristy of Toronto
10 King's College Road, Toronto, Ontario, M5S 3G4 Canada

Abstract

Touch-sensitive devices are becoming increasingly wide-spread, and consequently gestural interfaces have become familiar to the public. Despite the fact that many gestures require frequently dragging, pinching, spreading, and rotating the finger-tips, there currently does not exist a human performance model describing this interaction. In this paper, a novel user performance model is derived for virtual object manipulation on touch-sensitive displays, which involves simultaneous translation, rotation, and scaling of the object. Two controlled experiments with dual-finger unimanual manipulations were conducted to validate the new model. The results indicate that the model fits the experimental data well (with R^2 and R values above 0.9), and performs the best among several alternative models. Moreover, based on the analysis of the empirical data, the simultaneity nature of manipulation in the task is explored and several design implications are provided.

Keywords:
Modeling, human factors, multi-touch displays, virtual object manipulation, Fitts' law

1. Introduction

Recent advances in technology have made touch-sensitive displays affordable and widely available, ranging from large tabletops to small mobile devices such as tablets and cell phones. These advances have brought direct manipulation and multi-touch interaction to the general population for the first time. This has expanded the interface capabilities of modern computers and mobile computing devices, enabling interface designers to use a richer set of gestures, including not only gestures involving one or multiple fingers such as flicking, pinching and twisting, but also whole hand gestures (Cao et al., 2008b; Wigdor et al., 2011). Such direct manipulation based interactions have two main advantages (Nacenta et al., 2009; Wigdor et al., 2011): 1) they resemble real object manipulations in the physical world which lead to arguably more natural interactions, and 2) they allow users to perform operations on multiple degrees of freedom (e.g., translation and rotation) simultaneously which have the potential to increase the efficiency of complex manipulations.

Although gesture-based computer interfaces have been available for more than a decade, a comprehensive performance model has not yet been developed for them. Consequently, these techniques are not built upon a relatively solid theoretical foundation thus it is difficult to comparatively evaluate, model, or predict human performance for the latest generation of user interfaces. This contrasts with traditional pointing and dragging-based interactions where Fitts' law (Fitts, 1954; MacKenzie, 1992) can be used to analyze the performance of the mouse, stylus, or finger-tip. For example, Fitts' law has made it possible to evaluate and compare pointing devices (Soukoreff and MacKenzie, 2004), to improve the efficiency of user interfaces based on the prediction of movement times (Soukoreff and MacKenzie, 1995; MacKenzie and Soukoreff, 2002), and to create novel interaction techniques according to the optimization of Fitts' law (McGuffin and Balakrishnan, 2002). In short, movement models enable researchers to improve existing user interfaces, and to create novel interaction techniques. However, this sort of model has yet to be developed for touch-sensitive multi-degree of freedom interface technologies.

Our long-term goal is to develop a performance model for the range of multi-touch interactions that are emerging today, although in this paper we will focus on a common subset of manipulation gestures — unimanual dual-finger pinching, panning and twisting. As the first attempt to model multi-touch interaction, we selected this manipulation task because it is an example of a common everyday activity that users face when, for instance, adjusting the arrangement of photos in a photo album, sorting and organizing a number of documents on the tabletop, initiating specific commands via touch gestures, or playing cards during a game, where both the speed and the accuracy of the multi-touch gestures are concerned by designers. Zhao et al. (2011) describes results based on limited preliminary data from our initial exploration of this problem. Specifically, our objective is to construct a mathematical model of multi-touch manipulation that accommodates the gestures pertaining to the translation, rotation, and scaling of 2D virtual objects. In a similar vein to Fitts' law, our model shall relate

*Corresponding author. Tel.: +1 647 818 1919
Email addresses: jianzhao@dgp.toronto.edu (Jian Zhao), will@dgp.toronto.edu (R. William Soukoreff), ravin@dgp.toronto.edu (Ravin Balakrishnan)

the time, accuracy and geometry of a multi-touch manipulation task. But we deal with more complicated movements, which involve more degrees of freedom (i.e., x and y position, orientation and scale ratio) and heterogeneous actions of manipulating an object (i.e., translation, rotation and scaling). Thus in this work, we are interested in successful multi-touch docking tasks with acceptable tolerance (accuracy), and we present our model as a generalization of Fitts' law to higher dimensional tasks.

The paper is organized as follows. After a review of related work, we describe the two key problems of coupling the distance metric with Fitts' law to model multi-touch manipulation tasks — the non-linearity of scaling and the diversity of measurement units. Next we will propose our new model as an extension of Fitts' law, but including solutions for the two problems. Then, because of many factors involved in this complicated multi-touch manipulation task, we describe two consecutive experiments investigating different aspects of the task and provide empirical data to validate the new model; the regression analysis indicates that our model accounts for the empirical data with R and R^2 values above 0.9. We also report our experimental findings on users' simultaneity of the control over multiple movement components during the task. Finally, we conduct discussions about the proposed model including its generalization to Fitts' law and comparisons to alternative models.

2. Related Work

This section reviews two main related areas of the previous work — theoretical movement models of user performance and multi-touch manipulation techniques.

2.1. Movement Modeling

The preeminent movement model in human-computer interaction is Fitts' law, which has been used to model human performance at operating common pointing devices. Fitts' law defines the difficulty ID and movement time MT of a rapid aimed pointing task in terms of the distance D and a target width W,

$$MT = a + b \cdot ID, \qquad ID = \log_2\left(\frac{D}{W} + 1\right). \tag{1}$$

There exists extensive published research that extends the original 1D Fitts' law pointing task (see Eq.(1)) to multi-dimensional scenarios. MacKenzie and Buxton (1992) proposed several different formulas for 2D tasks for pointing at a rectangular target. Two formulations of the index of difficulty were found to highly correlate with experimental data: the first one reduces the 2D task to a 1D task by considering the target width W' to be the constraint of the target in the direction of movement, and the second one uses $\min(W, H)$ as the Fitts' law width where W and H are the width and height of the target. Accot and Zhai (2003) accommodated the angle of approach by proposing a weighted Euclidean model with the formulation,

$$ID_{WtEuc} = \log_2\left(\sqrt{\left(\frac{D}{W}\right)^2 + \eta\left(\frac{D}{H}\right)^2} + 1\right). \tag{2}$$

A more recent study by Grossman and Balakrishnan (2005) employed a probabilistic approach that can be generalized to 2D targets with any shapes. Also, various models of 3D target pointing have been explored by extending the ideas of computing 2D pointing task difficulty (Grossman and Balakrishnan, 2004).

In addition to translational movements, early studies indicated that Fitts' law can model rotary tasks (such as dial-turning) (Knight and Dagnall, 1967) yielding performance measures (i.e., throughput values) that are similar to those observed in translational movements (Crossman and Goodeve, 1983). Recently, Stoelen and Akin (2010) investigated real objects manipulations involving clockwise rotation and translation, proposing a model where the task difficulty was taken to be the sum of the difficulties of the translation and rotation components, which were in turn defined by expressions that are identical to the 1D Fitts' law index of difficulty.

Movement models for target acquisition in multi-scale interfaces have also been explored in several studies. Hinckley et al. (2002) found that Fitts' law can be used to model aimed scrolling interactions. On the other hand, when the target is not known ahead, Andersen (2005) proposed a simple linear model of such scrolling tasks, based on the notion of constant maximum scrolling speed. Some other models have also been proposed to model the acquisition of dynamically revealing targets with different environments, such as hand-held devices (Cao et al., 2008a) and multi-touch displays (Zhao et al., 2011). McGuffin and Balakrishnan (2002) investigated the application of Fitts' law in acquisition tasks of expanding targets (e.g., clicking items on the Mac OS dock panel), in which the index of difficulty is computed from the expanded target width. For pointing tasks in multi-scale electronic worlds with panning and zooming interactions, Guiard and Beaudouin-lafon (2004) introduced the scale variable and proposed a model by applying Fitts' law with the "zoom distance" defined under their multi-scale pointing paradigm; experiments of bimanual interactions using styluses and joysticks are conducted to validate the model.

Although these publications each consider interaction involving a dimension other than simple translation, they are all limited to their specific domains, and so it is unclear how they may be generalized to tackle multiple degrees of freedom interaction consisting of a combination of translation, rotation, and scaling, each of which have their own distinct nature. Also, the published empirical data all pertains to indirect interaction (using either a mouse, joystick or track-point) rather than direct manipulation (using fingers directly on a touch-sensitive display), except a recent study conducted by Bi et al. (2013). However, unlike our work which models a continuous multi-touch manipulation process involving many factors, their model only explores discrete finger touches for target acquisitions on the screen.

2.2. Multi-Touch and Multi-Dimensional Manipulation

Multi-touch interaction research is new enough that descriptive models have not yet been published — the multi-touch literature is primarily focused on developing novel interaction gestures, and exploring the possibilities of multi-touch manipulations on multiple degrees of freedom.

An important advantage of techniques utilizing direct touch manipulation is that they improve both coordination and parallelism (Forlines et al., 2007). Studies have also shown that people are more efficient on tabletop interfaces with multi-touch gestures when doing tasks that requires direct manipulations in the physical world (such as object sorting) (North et al., 2009). Multi-touch techniques have been applied in many applications and proved to be efficient. Cao et al. (2008b) observe that the shape of the area of contact between the hand and the touch-sensitive surface provides additional information regarding gestures, that may be used to increase the interaction bandwidth. Magic desk (Bi et al., 2011) enriches the traditional mouse and keyboard desktop interactions by integrating multi-touch technologies. Although multi-touch gestures are generally intuitive, improving the efficiency with which users learn different kinds of gestures is important. Freeman et al. (2009) describe a visualization technique for portraying gestures, and include a taxonomy of multi-touch and whole-hand gestures.

In the context of multiple degrees of freedom manipulation, some researchers have quantitatively studied users' abilities of simultaneously manipulating multiple movement parameters of different tasks, such as object transportation and orientation on a 2D surface (Wang et al., 1998), 3D object docking tasks with 6 degrees of freedom (Zhai and Milgram, 1998; Masliah and Milgram, 2000), and virtual 3D object orientation (Veit et al., 2009). However, none of the above studies involve the scaling operation or multi-touch interfaces, which we are interested in here.

Moreover, some studies have investigated multi-touch interaction techniques that either deliberately integrate or separate the individual degrees of freedom. For example, the *Rotate'N Translate* technique (Kruger et al., 2005) combined the control of rotation and translation using only a single touch-point, providing better support of comprehension and communication. However, the ability of operating on more degrees of freedom with multi-touch may cause problems when users intend to perform only a subset of manipulations. Nacenta et al. (2009) proposed several interaction techniques that separate the spatial manipulations (i.e., translate, rotate and scale) in multi-touch interfaces by filtering or classifying the movements. Along similar lines, Wigdor et al. (2011) employed the hand shape information for setting constraints on the manipulatable degrees of freedom to separate the user interactions.

There have been several attempts to empirically study multi-touch gestures. Hoggan et al. (2013a,b) conducted experiments to examine factors (such as angle and position) that could affect performance and ergonomics of two-finger pinch and rotation gestures. Nguyen and Kipp (2014) specifically studied the orientation factor in translation-rotation multi-touch tasks, and showed that movement combinations in different directions are more tiresome. But those studies have not looked into multi-touch manipulations that include a combination of translation, rotation, and scaling, which is of interest in this paper. Instead of studying the actual multi-touch actions, researchers have investigated factors that affect how users plan initial grasps for

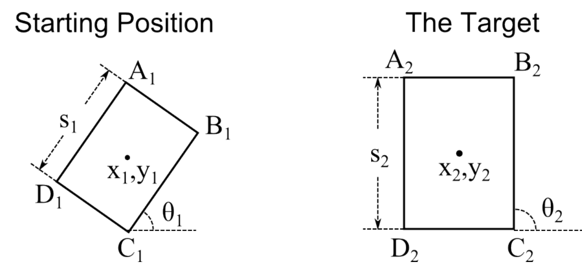


Figure 1: Multi-touch 2D object manipulation paradigm

translating and rotating virtual objects on a tabletop (Olafsdottir et al., 2014).

However, none of the above studies provide a basic theoretical understanding of the multi-touch manipulation. Thus it could be difficult to tell which gestures or techniques are best suited to specific applications. In contrast, a user performance model can not only provide guidelines to interface gesture designers but also implicitly describe the nature of the task, which offers a fundamental framework for previous studies.

3. A Model of Multi-touch Manipulation

Our objective is to formulate a human performance model analogous to Fitts' law but extended so as to encompass the more generalized multi-touch interaction of interest here, that includes translation, rotation, and scaling.

The physical placement of a 2D object has the following degrees of freedom: position x and y, orientation θ, and size scaling s. Thus the state of an object can be represented as a 4-tuple that includes these four quantities. In this framework, an object manipulation task is mathematically equivalent to the changing of the state of an object from its initial orientation vector $\vec{p}_1 = (x_1, y_1, \theta_1, s_1)$ to the result one $\vec{p}_2 = (x_2, y_2, \theta_2, s_2)$, as shown in Figure 1. Intuitively this leads us to want to define the "distance" of a multi-touch manipulation as the standard Euclidean distance between the start and end vectors $|p_1 - p_2|$, which we could then plug into the Fitts' law equation. Unfortunately, things are not quite so simple; as the following two sections will reveal, the naive Euclidean distance cannot be applied here — there are two problems that must be addressed first.

3.1. Non-linearity of Scaling

The first problem is that numerically the scale quantity does not behave linearly as translation and rotation do. The two actions of scaling an object to twice its size (so, for example, the scale goes from $s_1 = 1$ to $s_2 = 2$), and then returning it to its original size by halving its size (this time going from $s_1 = 1$ to $s_2 = 0.5$), both require approximately the same effort when achieved by the two-finger spreading gesture, followed by the two finger pinch. But the magnitude of the scale quantities suggests that it is more difficult to enlarge an object than to shrink it, because the difference of scales is larger for increasing the size ($\Delta s = s_2 - s_1 = 1$) than for decreasing the size ($\Delta s = -0.5$), which is not correct.

Another informative example of the nonlinearity of scale arises in the case of doubling the size of an object twice. Each individual doubling operation results in a change of scale of 1 as above ($s_1 = 1$ to $s_2 = 2$, $\Delta s = 1$), which suggests that together the two operations should yield a combined change of scale of 2 ($\Delta s_1 + \Delta s_2 = 1 + 1 = 2$). However, if analyzed together (as if they were a single operation), the two operations yield a different total change in scale ($s_1 = 1$ to $s_2 = 4$, $\Delta s = 3$, not 2). This differs from, for example, translation where the effect of moving an object 10 cm, followed by another 10 cm, is equivalent to a single movement of 20 cm.

3.2. Diversity of Quantity Units

If our task was to find a representative combination of two or three orthogonal measures of distance (e.g., distances in the Cartesian x, y, and z directions), where the distances are measured in meters, then we may expect that the result would be a quantity measured in meters, and that a movement in the x direction of 1 meter (i.e., $\Delta x = 1$, $\Delta y = \Delta z = 0$), would be equivalent (in terms of the total distance moved) to a movement in the y direction of 1 meter (i.e., $\Delta y = 1$, $\Delta x = \Delta z = 0$). But in our case the three separate quantities we wish to combine (position, rotation, and scale) each have different units and ranges. Simply put, one degree of rotation is not equivalent to one meter of movement. Resolving this incompatibility requires finding an equivalence between degrees of rotation, meters of distance, and units of scale, ideally such that the effort to achieve a certain level of accuracy in each dimension would be equal. Nevertheless, because the units and ranges of our four quantities are different we cannot simply apply a 4-dimensional standard Euclidean distance metric.

3.3. Model Derivation

We propose addressing the non-linearity of scaling by applying a logarithm transformation, which has been used in studies of modeling target acquisitions in multi-scale views (Guiard and Beaudouin-lafon, 2004). This approach achieves sensible values for the relative difficulty of adjusting the scale. For example, leaving an object's size unchanged ($s_1 = s_2$) results in a difficulty of $\log_2 s_2 - \log_2 s_1 = \log_2(s_2/s_1) = 0$, and for the first example provided above, doubling the size of an object ($\log_2 2 - \log_2 1 = 1$) and halving the size of an object ($\log_2 0.5 - \log_2 1 = -1$) again yield intuitive results. The second example given above, doubling the size of an object twice, is also resolved by the logarithmic transformation. Therefore, we redefine the object state vector to $\vec{p} = (x, y, \theta, \log_2 s)$.

3.3.1. Distance Metric

To solve the problem of the diversity of quantity units, we employ the *weighted Euclidean distance* which normalizes the individual constituents using different coefficients. The weighted Euclidean metric has shown effective in modeling 2D and 3D pointing tasks along with the Fitts' law (Accot and Zhai, 2003; Grossman and Balakrishnan, 2004), however, no study has explored applying such metric to measure the "distance" of tasks that contains different quantity units.

Applying the weighted Euclidean distance to the object state vectors yields,

$$A = \sqrt{(\Delta x^2 + \Delta y^2) + \alpha\Delta\theta^2 + \beta\Delta s^2}, \tag{3}$$

where $\Delta x = x_1 - x_2$, $\Delta y = y_1 - y_2$, $\Delta\theta = \theta_1 - \theta_2$ and $\Delta s = \log_2 s_1 - \log_2 s_2$, and the quantities α and β represent the weights of the respective components of distance, where we assumed that the translation components, x and y, have identical weight because of they having the same quantity unit. The values of these weight parameters, which will be determined empirically, standardize the units of other components into that of the translational components (i.e., in meters).

Further combining the first two items we obtain

$$A = \sqrt{\Delta d^2 + \alpha\Delta\theta^2 + \beta\Delta s^2}, \tag{4}$$

where $\Delta d = \sqrt{\Delta x^2 + \Delta y^2}$ corresponds to the distance between the center of the object at the starting and ending positions of the manipulation.

Therefore our proposed model is found by inserting the distance A into the Fitts' law formula for the index of difficulty (Eq.(1)), assuming that the task tolerance is governed by a parameter W similar to the Fitts' law width,

$$ID = \log_2(\frac{A}{W} + 1). \tag{5}$$

3.3.2. Task Tolerance

There are two ways of formalizing the tolerance of this multi-touch manipulation task. First, tolerance values of each movement components can be specified individually and combined in a similar format to our proposed metric of the task "distance",

$$W = \sqrt{\Delta D^2 + \alpha'\Delta\Theta^2 + \beta'\Delta S^2}, \tag{6}$$

where D, Θ and S are user-specified thresholds and α', β' are weight parameters. Ideally, we should use the same weight values as in Eq.(4) (i.e., $\alpha' = \alpha$ and $\beta' = \beta$). However, in practical, it is very difficult to assign reasonable values to the weights before any experiments and analysis, because features of touch devices are usually unknown and different from each other. On the other hand, allowing parameters in W makes the model more complicated for regression analysis. Also, setting proper individual thresholds is less practical, because the task difficulty of each component is hard to balance when treating a successful docking manipulation as a whole. For example, if we set a 5-pixel threshold for translation and a 5-degree threshold for rotation, we may wonder: is moving 1 pixel as difficult as rotating 1 degree?

Another approach is to measure the overall manipulation task tolerance as an integrated process, which is more appropriate in our study, since we are interested in modeling successful multi-touch docking tasks. For this we propose the *average corner distance*, which is defined as the mean value of distances between corresponding corners of the object. As shown in

Figure 1, the average corner distance between the two object states can be computed from $\frac{1}{4}(\overline{A_1A_2} + \overline{B_1B_2} + \overline{C_1C_2} + \overline{D_1D_2})$.

However, for an UI element with a more complicated shape, this corner points based measurement may not be the best to represent the overall shape variation. For example, a start-shaped object that has about the same bounding box size as an rectangle may result in a smaller average rotation error, since its interior corner points are closer to the center. A more general measurement for arbitrary-shaped objects, by extending the corner metric, would be an integration of distances between all the corresponding "points" covered in the area of the object,

$$\frac{1}{A}\int_A ||p_1 - p_2|| dA, \tag{7}$$

where A is the area of the element and p_2 is the corresponding point of $p_1 \in A$ after the movement.

3.3.3. The Model

Finally, presuming a linear relationship between difficulty and movement time (analogous to Fitts' law), our formula for the movement time of this multi-touch manipulation task is

$$MT = a + b \cdot \log_2(\frac{A}{W} + 1), \tag{8}$$

where A is the distance defined in Eq.(4) and the values of the weight parameters will be determined empirically; W is predefined task tolerance computed from one of the above measurements and in our paper we use the average corner distance for practical reasons.

4. Experiment 1

The multi-touch manipulation tasks involve many factors, including the distances and directions of the three variables — translation, rotation and scaling. It is impractical to do full-crossed study with many levels of all the factors. Thus the purpose of the first experiment was to explore all aspects of the task and gather empirical data to identify important conditions requiring further explorations and initially test the model described above.

4.1. Participants

Ten volunteers (6 males and 4 females, including 2 novice users) participated in the study. The participants were members of the university community (graduate students and researchers). All participants were right-handed. The participants' average physical measures were: age 23.8 ($SD = 4.0$), height 173 cm ($SD = 10$), hand length 17.7 cm ($SD = 1.8$), hand breath 11.2 cm ($SD = 1.5$), hand span 18.8 cm ($SD = 1.4$), index-finger-to-thumb span 15.1 cm ($SD = 1.8$) and forearm length 43.7 cm ($SD = 2.3$).

4.2. Apparatus

Participants were asked to manipulate 2D virtual objects on a Microsoft Surface tabletop v1.0. The display size was 30 inches (76 cm) diagonal, with a resolution of 1024×768 pixels. The Surface was physically raised by placing it on a stand 40 cm high, so the participants stood while operating it (this made it easier for the participants to reach the entire display, and avoided the problem of where the participants should put their knees while using the Microsoft Surface).

Custom experiment software was developed using the Microsoft Surface SDK. This software presented the manipulation tasks to the participants, and recorded the times that the participants took to perform the tasks. The implementation of the object manipulation algorithm followed the most intuitive and simplest approach: the translation distance equaled the movement of the center of the contact points; the degree of rotation was obtained from the angle difference of normalized vectors formed by two contact positions; the scale factor was computed by dividing the ending distance of two fingers with its starting distance; and no sticky effects were employed.

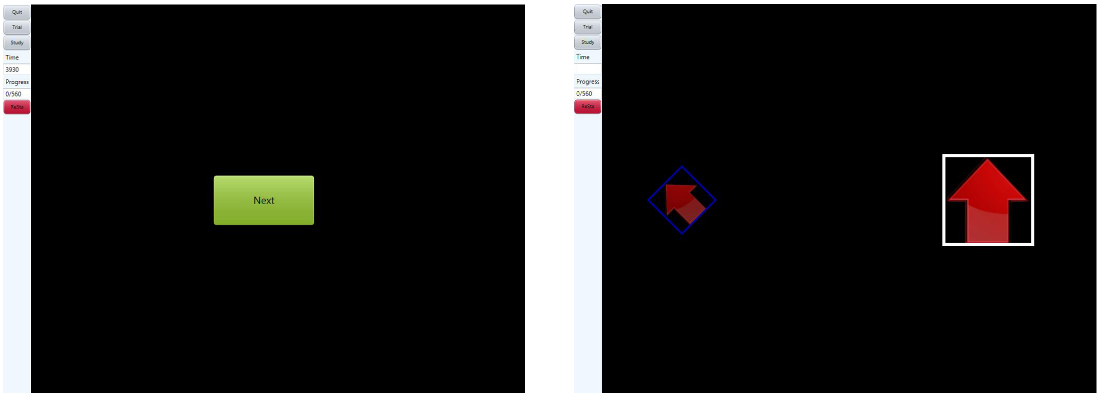
4.3. Procedure

The participants were observed as they completed a representative series of multi-touch interactions. The participants were asked to use the thumb and either the index finger or the middle finger (whichever was more comfortable) of their dominant hand to perform the manipulation tasks. We did not restrict the starting finger positions for realism to maximize the external validity. Clutching (lifting the hand off of the surface) during a manipulation was not allowed. Thus the manipulations had to be performed in a single rapid aimed motion, which followed the design of most Fitts' law studies because clutching reduces the fitting of the model (Casiez et al., 2008).

Before the study began, the experimenter demonstrated the manipulation task to the participant, and the participant performed 32 practice trials. The condition values were similar to, but not the same as, the values chosen for the study trials. Data from the practice trials was discarded.

Each trial began with the display of the green "Start" button (Figure 2a). After clicking the button, the manipulatable object (a blue bordered square) and target (a white bordered square) were presented (Figure 2b). The manipulatable object was translated, rotated, and scaled, in accordance with the condition being presented to the participant, the target squares were uniformly presented with 150 pixels width 150 pixels height, and 0 degrees of rotation. Both squares were filled with a red arrow background image to indicate their orientation (thus there was only one correct orientation). The manipulatable object was semi-transparent so that it did not occlude the target near the end of the trial as it overlapped it. Once the manipulatable object was close enough to the target that the average corner distance was less than a predefined threshold (20 pixels), the border of the object turned green to indicate that the trial was successful (i.e., not an error).

A trial was considered to be in error if the user failed to achieve the required accuracy (i.e., if manipulatable object was



(a) starting screen

(b) multi-touch manipulation task

Figure 2: Snapshots of the experimental software.

released when the border of the object was not green), or if the experiment requirements were violated (for example if more than 2 fingers were in contact with the display surface). Trials in which errors occurred were presented to the participants again until the participants succeeded.

Each participant took about an hour to complete the study, so to avoid fatigue the participants were encouraged to rest between blocks. Following the study a short informal interview was conducted to gather subjective feedback.

4.4. Design

The independent variables consisted of the three movement parameters: translation distance, rotation angle, and size scaling. A within-subject design was employed with the above conditions fully-crossed. We also experimented with possible direction-of-movement effects of the three parameters, specifically, eight directions of the translations; clockwise and counter-clockwise rotations; and pinching-in and pinching-out scalings. In particular, the levels of conditions were: translation distance: 0, 250, 500 pixels; translation direction: N, S, E, W, NE, NW, SE, SW; rotation: -30, 0, 30 degrees; and scale ratio[1]: $2^{-0.5}$, 1.0 (2^0), $2^{0.5}$; totaling 216 movement conditions. Note that these conditions were fully crossed, and so for example, conditions involving only translation and scaling (because rotation = 0), or just translation (because scale = 1 and rotation = 0) were included, including the degenerate case where the subject only had to tap the display because no manipulation was needed (translation = 0, scale = 1 and rotation = 0).

We chose not to vary the task tolerance W because: 1) our focus here is to model successful docking operations as the first step, and 2) adding more levels of W will make the experiment contain too many conditions. Thus, we conducted a pilot study with 4 volunteers in order to explore the reasonable tolerance of a successful docking. In the end, the average corner distance with 20 pixels threshold subjectively felt the best to the experimenter and pilot subjects.

During the real study, each participant conducted 2 blocks of trials. Within each block, the 216 trials were presented to participants in a random order. In summary, the experiment consisted of 10 participants × 2 blocks × 3 translations × 3 scales × 3 rotations × 8 translation directions = 4320 trials.

The dependent variables were movement time and error rate. The movement time was the time that elapsed between when they first touched the display surface (beginning the manipulation) and when they lifted their fingers (denoting the end of the trial). Other quantities that were also recorded by the software include the number of errors (i.e., retrials), the finger contact positions, and a timestamped log of the location, angle, and scale of the manipulatable object throughout the manipulation operation.

5. Results and Analysis of Experiment 1

5.1. Movement Time

The average movement time MT of this multi-touch manipulation task was 1691 ms ($SD = 737$). Repeated measure ANOVAs indicated significant effects for translation ($F_{2,18} = 223.9, p < .001$), scale ($F_{2,18} = 25.95, p < .001$), and rotation ($F_{2,18} = 24.61, p < .001$) on MT. Significant interactions were found in translation×scale ($F_{4,72} = 12.02, p < .001$), translation×rotation ($F_{4,72} = 17.71, p < .001$), and rotation×scale ($F_{4,72} = 8.293, p < .001$).

The statistics of movement times on all the movement factors (including translation directions) are shown in Figure 3. By comparing the mean movement times of each level of the conditions, we identified several interesting facts. First, the movement time increased monotonically as the absolute traveling distance for each of the three movement parameters increased. Second, rotating clockwisely and counter-clockwisely produced similar average times to one another (1807 ms for -30 degrees and 1770 ms for 30 degrees), and similar phenomena were observed in scaling operations (1791 ms for scale ratio $2^{-0.5}$ and 1823 ms for $2^{0.5}$). Third, translation directions had some effects on the movement time ($F_{7,63} = 4.970, p = .020$);

[1]Note that the scale ratio levels were chosen symmetrically in the log-transformed space defined in previous sections.

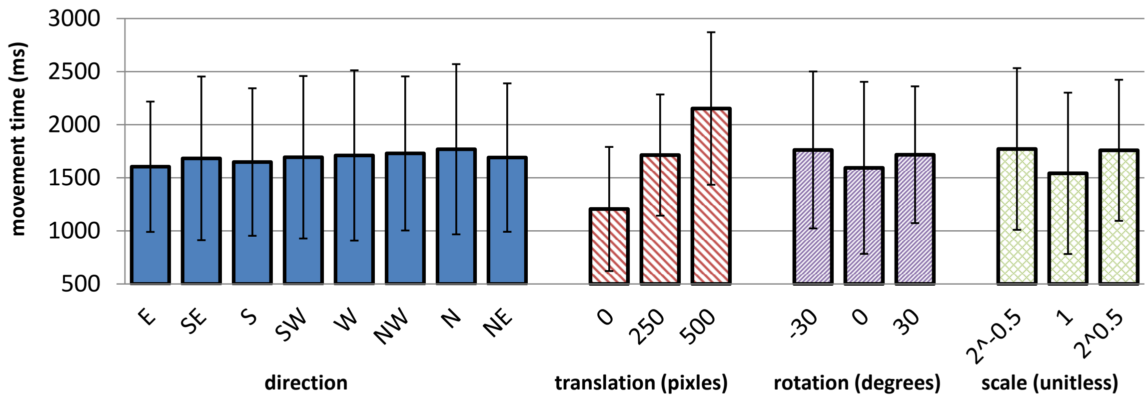


Figure 3: Effects of different factors on the movement time in Experiment 1.

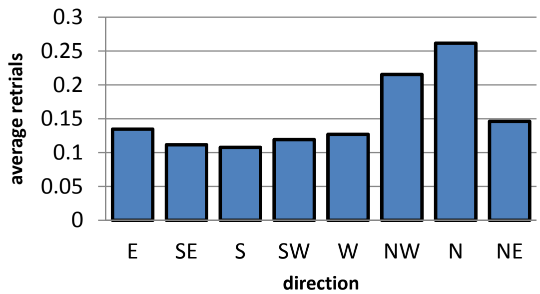


Figure 4: Average number of retrials on all translation directions in Experiment 1.

more specifically, the times were a little smaller in near peripersonal space (direction E, SE and S) and larger in far peripersonal space (direction W, NW and N), from a right-handed perspective.

5.2. Errors: Numbers of Retrials

Errors occurred when participants failed to manipulate the object to the targeted position to within the predefined tolerance. For this analysis, errors that were caused by violating the experiment requirements (e.g., touching the display surface with more than 2 fingers) were removed. The overall average error rate was 7.45% (322 out of 4320 trials). There was a significant effect on the average number of retrials for translation ($F_{2,18} = 5.191, p = .016$) and no significant effects were found for rotation or scale. Movement directions also slightly influenced the errors ($F_{7,63} = 2.709, p = .016$). As shown in Figure 4, more errors were encountered when manipulating objects away from the body (direction NW, N and NE) than that when manipulating objects towards the body (direction SE, S and SW), of which the North direction produced the most number of retrials (70 retrials) and the South direction was the least (29 retrials).

5.3. Accuracy: Average Corner Distance

We also recorded the ending corner distances of all the trials, which measured the tolerance of the multi-touch manipulation task. The mean value of corner distances of all the trials was 8.7 pixels ($SD = 4.2$). Figure 5 shows the statistics of all the factors, revealing similar patterns to effects of different factors on the movement time. ANOVA tests indicated that there were significant effects on translation ($F_{2,18} = 10.69, p < .001$), scale ($F_{2,18} = 16.75, p < .001$) and direction ($F_{2,18} = 3.406, p = .004$). Similar to movement times and errors, directional effects were observed, where directions in far right-handed peripersonal space (such as NW and N) had relative larger corner distances (less accurate) in general.

5.4. Model Fitting

We fitted the multi-touch manipulation model appearing in Eq.(8) to our experimental data (non-error trials only) using the MATLAB function `fminsearch`[2] that is commonly used in nonlinear unconstrained optimization. It attempts to find a minimum of a multivariate scalar function with the simplex search method, starting at an initial estimate. The objective function was the sum of squared errors between predicted and measured movement times and we used random values as the initial estimate. The data points fitted were movement times of the 216 movement conditions described the previous section averaging from values of 10 participants. The results of the regression analysis were: $a = 361$ ms, $b = 369$ ms/bit, $\alpha = 7.02$ pixel^2/degree^2, and $\beta = 2.20 \times 10^4$ pixel^2. As Figure 6 shows, our model fits the empirical data very well, yielding R^2 value 0.90 ($R = 0.94$).

It is worthy to note that we did not average the movement times across the directions of translation (N, S, E, W, NE, NW, SE, and SW), scale (pinching-in and pinching out) and rotation (clockwise and counter-clockwise), although the model in Eq.(8) does not have parameters to describe these directional characteristics.

Usually in Fitts' law studies, where participants often point to two targets back and forth, model fitting analysis considers only factors appearing in the formulation by taking averages of movement times on the direction factors, such as Accot and

[2] http://www.mathworks.com/help/toolbox/optim/ug/fminsearch.html

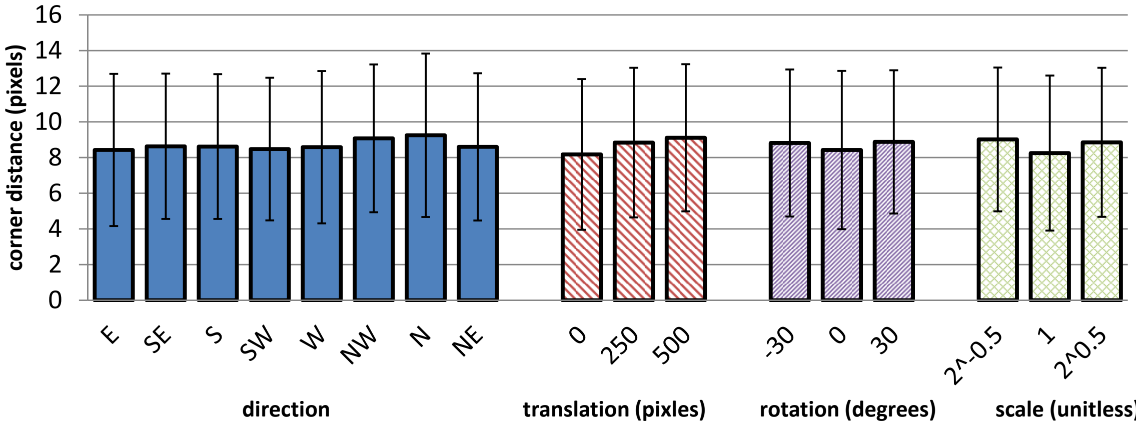


Figure 5: Effects of different factors on the average corner distance in Experiment 1.

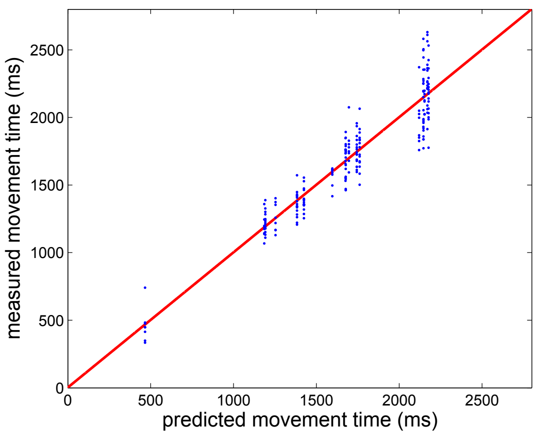


Figure 6: Model fitting with the empirical data of Experiment 1 (R^2 = 0.90, R = 0.94).

Zhai (2003) and Grossman and Balakrishnan (2004). Since only a few levels of each factor had been chosen in this experiment, we noted that it would significantly reduce the number of data points by averaging them across movement directions, which would also increase R^2 values. Therefore the fitting results above indicated that the performance of our model could be even better.

6. Experiment 2

The results of the above experiment indicates that our model can describe the empirical data very well. However, only a few levels of each movement factor were experimented since the eight translation directions could easily explode the study. The goal of the second experiment was to explore more levels and larger ranges of the three movement factors appearing in Eq.(8) in order to further validate the model. For practical issues, we chose two directions in which participants had the extreme performances — North (that had the slowest movement time: 1769 ms) and East (that had the fastest movement time: 1604 ms), which also happened to have the extreme variances of movement times. Similar experimental design approach has been used by Banovic et al. (2011) to experiment multi-touch marking menu selections. Thus we believe that a model which can explain user performance on these two extreme directions is very likely valid for all other directions. The apparatus and procedure of Experiment 2 were similar to those in Experiment 1.

6.1. Participants

Another group of 10 participants, (6 males and 4 females; all right-handed) was recruited for the study. They were university graduate students and two of them were novice to touch-sensitive devices. The participants' average physical measures were: age 24.3 (SD = 1.3), height 169 cm (SD = 6.8), hand length 16.4 cm (SD = 1.6), hand breath 11.1 cm (SD = 1.4), hand span 19.2 cm (SD = 1.5), index-finger-to-thumb span 14.4 cm (SD = 1.2) and elbow length 41.8 cm (SD = 2.5).

6.2. Design

The independent variables, i.e., the three movement parameters including translation, scale and rotation, were the same as those in Experiment 1. A similar within-subject design was employed, in which the three variables were fully-crossed and their direction-of-movement effects were included. More specifically, the levels of conditions were: translation distance: 0, 200, 400, 600 pixels; translation direction: N, E; rotation: -135, -90, -45, 0, 45, 90, 135 degrees; and scale ratio: $2^{-0.9}$, $2^{-0.45}$, 1.0 (2^0), $2^{0.45}$, $2^{0.9}$; totaling 280 movement conditions. The same to Experiment 1, 2 blocks of the 280 trials were presented to each participant in a random order. Therefore, the whole experiment consisted of 10 participants × 2 blocks × 4 translations × 5 scales × 7 rotations × 2 translation directions = 5600 trials. The dependent variables we measured were the same as the previous experiment.

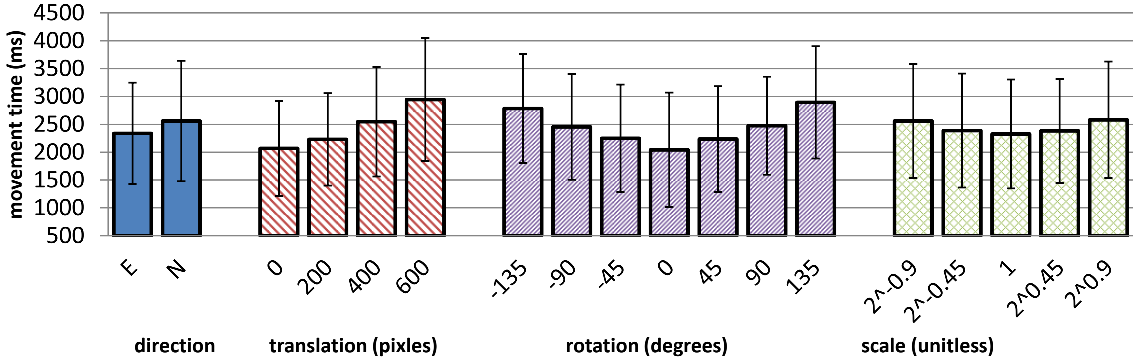


Figure 7: Effects of different factors on the movement time in Experiment 2.

7. Results and Analysis of Experiment 2

7.1. Movement Time

We conducted similar analysis of the movement time as in Experiment 1. The empirical data showed the average movement time of the multi-touch manipulation task was 2448 ms ($SD = 1006$). Significant effects were found on all the movement factors: translation ($F_{3,27} = 81.53, p < .001$), scale ($F_{4,36} = 13.61, p < .001$), and rotation ($F_{6,54} = 65.31, p < .001$). Also, repeated measure two-way ANOVAs indicated that there were significant interactions on translation×rotation ($F_{12,171} = 5.133, p < .001$) and rotation×scale ($F_{24,306} = 1.910, p = .008$).

The effects of the three movement factors and the translation direction on the movement time are shown in Figure 7. Concurred with our findings in Experiment 1, the movement time increased as the absolute traveling distance of each factor increased. Also, the corresponding movement times of two symmetrical traveling directions (i.e., clockwise and counter-clockwise rotations; pinching-in and pinching-out scales) appeared to be similar. There was also a significant effect for translation direction ($F_{1,9} = 20.04, p = .002$).

7.2. Errors and Accuracy

After conducting the same preprocessing of the log files as in Experiment 1, we observed that the overall error rate was 12.5% (703 retrials out of 5600 trials), which was greater than Experiment 1 since the difficulty of tasks increased. Significant effects of the number of retrials were found on factor translation ($F_{3,27} = 4.159, p = .015$) and rotation ($F_{6,54} = 2.322, p = .045$). Concurred with the first experiment, the error rates increased as the translation distance increased. However, in general counter-clockwise rotations showed more errors (14.0%) than clockwise rotations (10.8%), of which rotating at -135 degrees was the largest (15.7%). This unsymmetrical results may be because right-handed participants likely started the manipulation by positioning their two fingers in the posture that was difficult to rotate large angles counter-clockwisely.

With similar analysis of the task tolerance as in Experiment 1, ANOVAs indicated that there were significant effects for the tolerances (ending corner distances) of the manipulation tasks on translation ($F_{3,27} = 7.067, p = .001$), scale ($F_{4,36} = 3.082, p = .028$), and rotation ($F_{6,54} = 5.737, p = .001$). The same patterns of the distributions of corner distances with respect to the three movement parameters (as in Figure 5 of Experiment 1) were observed in the second experiment, which confirms such patterns in lager value ranges of the movement factors.

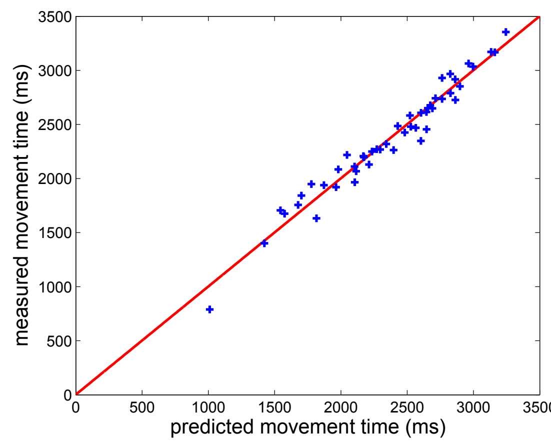
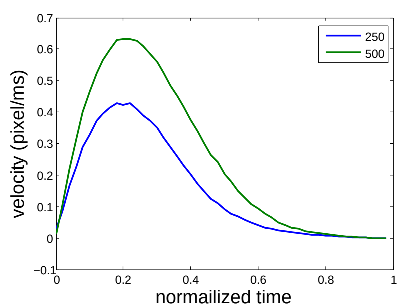


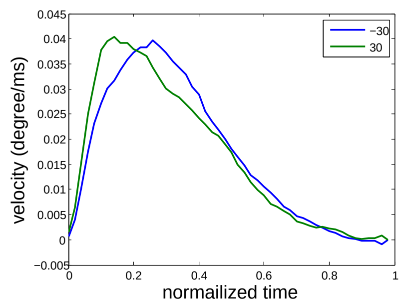
Figure 8: Model fitting with the empirical data of Experiment 2 ($R^2 = 0.91, R = 0.95$).

7.3. Model Fitting

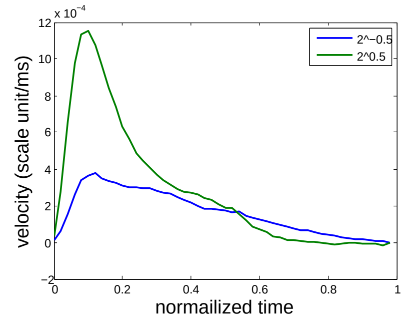
We conducted regression analysis with the empirical data collected in the second experiment in a similar way, i.e., using the same MATLAB fitting method with the same object function. The fitting results are shown in Figure 8. The data points were averaged movement times of 48 movement conditions with different traveling distances (4 translations×3 scales×4 rotations). The fitting yielded the R^2 value 0.91 ($R = 0.95$) with the estimated movement parameters: $a = 216$ ms, $b = 494$ ms/bit, $\alpha = 13.4$ pixel^2/degree^2, and $\beta = 2.42 \times 10^4$ pixel^2, indicating that our model was still valid with the experimental data having more levels and larger value ranges of the movement parameters.



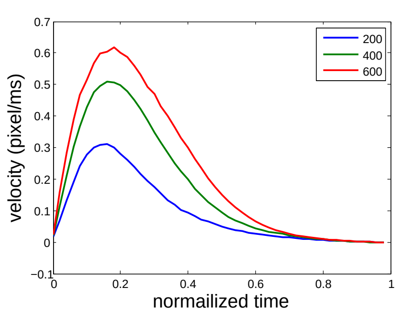
(a) translation (Experiment 1)



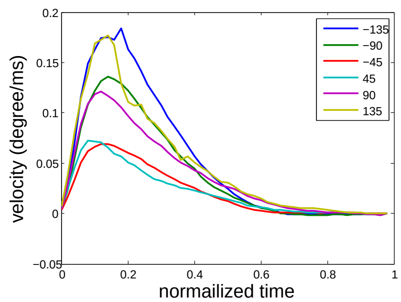
(b) rotation (Experiment 1)



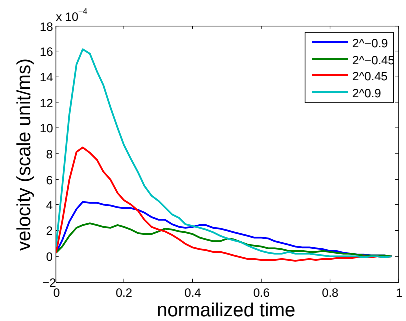
(c) scale (Experiment 1)



(d) translation (Experiment 2)



(e) rotation (Experiment 2)



(f) scale (Experiment 2)

Figure 9: Average velocity profiles of different conditions for the three movement parameters; the time domain of these velocity curves was normalized before the aggregation.

8. Results and Analysis on Movement Simultaneity

To better understand the characteristics of this multi-touch manipulation task, we investigated the velocity profiles of different movement conditions in this multi-touch manipulation task to study if and how users could simultaneously control multiple degrees of freedom (i.e., translation, rotation and scaling) during the experiments. As Figure 9 indicates, by analyzing the timestamped logs generated in our experiments, we normalized and aligned these velocity profiles in their time domains (note that the identity conditions, translation = 0, scale = 1, and rotation = 0, were removed from this data).

8.1. Movement Parameters on the Shapes of Velocity Profiles

It is known that velocity profiles of rapid aimed pointing tasks are typically bell-shaped, and distance and target width have systemic effects on the shape of the curves; larger distances result in higher peak velocities, and narrower target widths skew velocity profiles to the right (MacKenzie et al., 1987). Figure 9a and Figure 9d show that the height of the translational velocity peak increases with the distance, which is similar to that of translation-only Fitts' law pointing tasks. The velocity profiles of the rotational and scaling components, although express similar relationships between the velocity peak and movement amplitudes (i.e., absolute rotational degrees and scaling ratios), however, fluctuate more, resulting in multiple speed plateaus and small peaks. Particularly, Figure 9c and Figure 9f reveal that the participants demonstrated the least control over scaling, and that they sometimes overshot their scaling target, necessitating corrections (viz., negative velocities near the end of the manipulation task). Further, rotational velocity profiles of clockwise and counter-clockwise movements have similar curve shapes (Figure 9b and Figure 9e), indicating that the rotational direction seems to impose little effect on the shape of the velocity profile.

If viewing these sub-figures in Figure 9 altogether, we could observe that approximately at timestamp 0.18, where most of the translational and rotational velocities reached the peak velocities, most of the scaling velocity profiles already passed the highest peak values (which were around timestamp 0.1 approximately), indicating that participants could make the adjustments of distances and angles simultaneously along the whole manipulation process, but that they tended to scale the objects at the beginning of the task to have a comfortable launch-on finger distance, and then to make fine adjustments later.

8.2. Simultaneity Measurement

To quantify the simultaneity of the control over the three components in this multi-touch manipulation task, the *M-metric* measurement (Masliah and Milgram, 2000) was computed for all the trials of both experiments. This metric is the product of both simultaneity and efficiency of a manipulating trajectory. The simultaneity of control equals the overlapping area of normalized error reduced curve for each component, and the manipulation efficiency is related to weighted average of the ratios of the length of the optimal trajectories. ANOVAs on the measurement values indicated: for Experiment 1, there were significant effects of translation ($F_{1,9} = 138.9, p < .001$), rotation ($F_{1,9} = 171.2, p < .001$), and scale ($F_{1,9} = 41.63, p < .001$); and for Experiment 2, translation ($F_{2,18} = 49.19, p < .001$), rotation ($F_{5,45} = 53.95, p < .001$), and scale ($F_{3,27} = 65.35, p < .001$). Significant interactions were also found on: translation×rotation ($F_{1,9} = 6.820, p = .028$), rotation×scale ($F_{1,9} = 73.34, p < .001$), and direction×rotation ($F_{7,63} = 2.693, p = .016$), for Experiment 1; and translation×rotation ($F_{10,90} = 10.64, p < .001$), rotation×scale ($F_{15,135} = 4.849, p < .001$), and direction×scale ($F_{3,27} = 3.264, p = .037$), for Experiment 2. Further pairwise mean comparisons indicated that participants had more simultaneity of the control when the translational distance was small, when the scale ratio was large, and when rotational direction was counter-clockwise (which may be due to the recruitment of only right-handed participants).

9. Discussion

In this section, we discuss and explore many aspects related to the research problem in order to further understand the nature of the multi-touch manipulation task and the proposed model.

9.1. Measurement of Accuracy

For the two experiments described in previous sections, in order to define a successful docking and express the multi-touch manipulation task tolerance, we employed the average corner distance rule derived from the pilot study, because the metric in Eq.(6) contained parameters that were best determined after the experiments. Such tolerance metric is less practical than the average corner distance but more precise in comparing different manipulations on the same kinds of devices when parameters are known through pilot studies.

To test if the average corner distance was a good measurement, we analyzed the relationships between the ending corner distances of all trials and their corresponding distance metric values computed from the final statuses of the objects and the targets, using the empirical parameters α, β and γ estimated from the experiments. Regression analysis indicated that there were high linear correlations ($R = 0.94$ for Experiment 1 and $R = 0.90$ for Experiment 2) between those two metrics (Figure 10). Since our model with the proposed distance metric was already validated by the experimental data, the linear correlations not only justified our usages of average corner distance metric in the experiments but also indicated that such metric could be used as a measurement of the accuracy of multi-touch manipulations in practice. However, further study is needed to explore the effects of different task tolerances on the object manipulations.

9.2. Relations to Fitts' Law

The model presented in this paper is a generalization of Fitts' law to a higher degrees of freedom task. It can be demonstrated that Fitts' law arises as a special case of our model by considering what happens when a movement task consists solely of translation, with no rotation or scale components, i.e., when $\Delta\theta \equiv 0$ and $\Delta s \equiv 0$. Therefore the model becomes,

$$MT = a + b \log_2 \left(\frac{|\Delta d|}{W} + 1 \right), \tag{9}$$

which is the formula for Fitts' law. Similarly, our model suggests that Fitts' law can also be applied to rotational only and scaling only tasks, for example, when $\Delta d \equiv 0$ and $\Delta s \equiv 0$,

$$MT = a + b \log_2 \left(\frac{|\Delta\theta|}{W \cdot \sqrt{\frac{1}{\alpha}}} + 1 \right). \tag{10}$$

Several previous studies have indicated that Fitts' law models rotational tasks well (Stoelen and Akin, 2010; Knight and Dagnall, 1967; Crossman and Goodeve, 1983). Guiard and Beaudouin-lafon (2004) used the zoom index, which has similar formulation as ours for transforming the scale component (i.e., $\Delta s = \log_2 s_1 - \log_2 s_2$), to model pointing tasks in multi-scale navigation.

However, object scaling-only tasks have not been tested by Fitts' law. With the scaling distance calculated after log transformations, we fitted the Fitts' law formula to the empirical data from the scaling only conditions of the second experiment (the value of β in the estimated parameters of the model in Experiment 2 was used to compute the target width). The model was found to fit the data reasonably well, $R^2 = 0.95$. Nevertheless, because we only had 5 levels for the scaling conditions, these results should be considered preliminary, further exploration of modeling scaling tasks with Fitts' law is left for future work.

9.3. Alternative Models

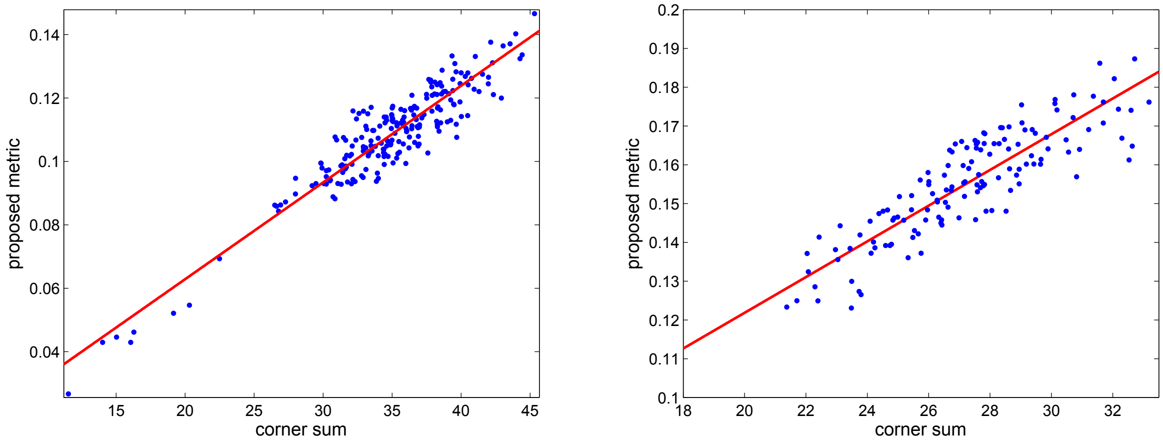
In the proposed model, we applied the weighted Euclidean distance along with the Fitts' law, allowing the incorporation of translation, rotation and scaling movement components to the multi-touch object manipulation. Of particular interest, we explored an alternative formulation (applying weighted Manhattan distance) of the index of difficulty for this task,

$$MT = a + b \cdot ID, \tag{11}$$

$$ID = \log_2 \left(\frac{|\Delta d| + \alpha|\Delta\theta| + \beta|\Delta s|}{W} + 1 \right), \tag{12}$$

which employs the weighted Manhattan distance instead. The fitting results of this model with our empirical data were: $a = 426$ ms, $b = 334$ ms/bit, $\alpha = 2.19$ pixel/degree, $\beta = 130$ pixel, and $R^2 = 0.83$ for Experiment 1; and $a = 309$ ms, $b = 438$ bit/ms, $\alpha = 3.67$ pixel/degree, $\beta = 147$ pixel, and $R^2 = 0.81$ for Experiment 2. Therefore, this new model has worse performance than our original model, indicating that the weighted Euclidean distance is more appropriate for this multi-touch manipulation task.

Moreover, we examined the fit of our empirical data with a model similar to the one proposed in Stoelen and Akin (2010), in which the task difficulty is a weighted sum of individual



(a) experiment 1 (R = 0.94)

(b) experiment 2 (R = 0.90)

Figure 10: The relationships between recorded final average corner distances and post-computed task "distance" values with the metric in Eq.(4).

index of difficulties of all three components. In this model, the movement time has the following formulation,

$$MT = a + l\log_2(\frac{|\Delta d|}{D} + 1) + m\log_2(\frac{|\Delta \theta|}{\Theta} + 1) + n\log_2(\frac{|\Delta s|}{S} + 1), \quad (13)$$

where a, l, m, and n are empirically determined coefficients, D, S, and Θ are tolerance parameters like the width W in Fitts' law, and Δd, $\Delta\theta$, and Δs have the same expressions as those in Eq.(4). Since in the experiments, where we are interested in successful multi-touch docking tasks, an overall tolerance threshold (i.e., the average corner distance) was used rather than setting individual tolerances for each components, thus in order to fit this model, we inserted the post-computed *effective widths* — D, S, and Θ — into the equation. The effective width, which is an adjustment for the tolerance term W in Fitts' law based on the spread of movement endpoints, is defined as $W_e = \sqrt{2\pi e}\sigma \approx 4.133\sigma$, where σ is the standard deviation of the movement endpoints (MacKenzie, 1992; Soukoreff and MacKenzie, 2004). And those studies have shown that applying the effective width to Fitts' law could improve the model performance in traditional pointing tasks. However, the fitting results with our experimental data indicated that this model still had worse performance than our proposed model: $R^2 = 0.79$ for Experiment 1 and $R^2 = 0.77$ for Experiment 2. Details of the fitting results of all the models in this paper can be found in Table 1.

In addition, alternative models can be created by using automated methods, as proposed by Oulasvirta (2014). Such techniques algorithmically generate equations according to certain constraints, which could quickly examine a large model space for a particular problem. However, the generated models, although can achieve high R^2 values for regression analysis, may have limited external validity, because they can be easily over-fitting the data. As opposed to traditional model development approaches that leverage observations and intuitions from real-world interactions, auto-generated expressions could be very complicated and have little physical meaning to explain the fundamental nature of user tasks. Nevertheless, it is an interesting future work for this paper to see what other models may come out from the auto-searching techniques and how similar they are to our models.

9.4. Implications on User Interface Design

The results of our study can be used to guide the development of user interfaces involving multi-touch manipulation and gestures from both design and theoretical perspectives.

First, our experimental results, including the movement times, task retrials and object docking accuracies, indicate that the task performance was more efficient in the near peripersonal space (i.e., direction E, SE and S; moving towards the user) than in the far peripersonal space (i.e., direction W, NW and N; moving away from the user), from a right-handed perspective (Figure 3, 4, and 5), which is aligned with some of the findings in previous studies (Mason and Bryden, 2007; Banovic et al., 2011; Hoggan et al., 2013a). This suggests that multi-touch interface designers should place more commonly-used targets or gesture commands in the near peripersonal space for table-top applications.

Moreover, although there was no significant effect of task retrial numbers on the rotation factor in Experiment 1, this became significant in Experiment 2 when the rotation range tested was larger. More specifically, counter-clockwise rotations resulted in more errors, which may be due to the fact that we only recruited right-handed users. Participants tended to place their fingers in the initial posture that was more easier to rotate large angles clockwisely, i.e., with the index-finger at the top-left direction of the thumb. This paralells the results of Olafsdottir et al. (2014)'s experiments, where they found that the initial grasp orientation for rotation tasks is largely affected by the starting and targeting configurations. Therefore, designers should consider avoiding multi-touch gestures with large rotation angles to balance the user performance for all cases.

Weighted Euclidean	$MT = a + b\log_2(\frac{1}{W}\sqrt{\Delta d^2 + \alpha\Delta\theta^2 + \beta\Delta s^2} + 1)$				
	a ms	b ms/bit	α pixel^2/degree^2	β pixel^2	R^2
Experiment 1	361	369	7.02	2.20×10^4	0.90
Experiment 2	216	494	13.4	2.42×10^4	0.91
Weighted Manhattan	$MT = a + b\log_2(\frac{1}{W}(\lvert\Delta d\rvert + \alpha\lvert\Delta\theta\rvert + \beta\lvert\Delta s\rvert) + 1)$				
	a ms	b ms/bit	α pixel/degree	β pixel	R^2
Experiment 1	426	334	2.19	130	0.83
Experiment 2	309	438	3.67	147	0.81
Stoelen and Akin (2010)'s	$MT = a + l\log_2(\frac{\lvert\Delta d\rvert}{D} + 1) + m\log_2(\frac{\lvert\Delta\theta\rvert}{\Theta} + 1) + n\log_2(\frac{\lvert\Delta s\rvert}{S} + 1)$				
	a ms	l ms/bit	m ms/bit	n ms/bit	R^2
Experiment 1	930	255	159	378	0.79
	effective widths: $D = 11.9$ pixel, $\Theta = 8.19$ degree, $S = 0.267$				
Experiment 2	1265	219	322	292	0.77
	effective widths: $D = 11.9$ pixel, $\Theta = 7.50$ degree, $S = 0.238$				

Table 1: Coefficients of all model fittings with the empirical data.

However, further studies should be conducted to systemically examine factors that influence a user's initial posture for the multi-touch manipulation tasks.

Last but not least, the simultaneity analysis of this multi-touch manipulation task indicates that users can well incorporate rotation and scaling operations into the execution of translational movements. Thus, in the case that this translation, rotation and scaling manipulation is transformed to a sequential and serial process, the total task completion time may increase significantly. If the main goal is to achieve fluid and natural interactions, gestures of a multi-touch interface should be designed in this concurrent way. However, participants demonstrated different control abilities over various movement components, where translation was the best and scaling was the least. So suppose that a virtual controller needs to be designed for adjusting multiple values with this multiple degrees of freedom manipulation, more important variables that require more precisions should be mapped to translational or rotational movements but not the scaling operation.

9.5. Bimanual Manipulation?

As the first attempt of studying the foundations of the simultaneous multi-touch manipulation task, we have developed a novel performance model and validated this model through two experiments with dual-finger unimanual manipulations. We did not choose to study bimanual conditions in our experiments because: 1) unimanual manipulations are more commonly used across devices with various display sizes and different usage scenarios, and 2) investigating too many factors in a single experiment would further explode our study where the fatigue of participants could affect the results.

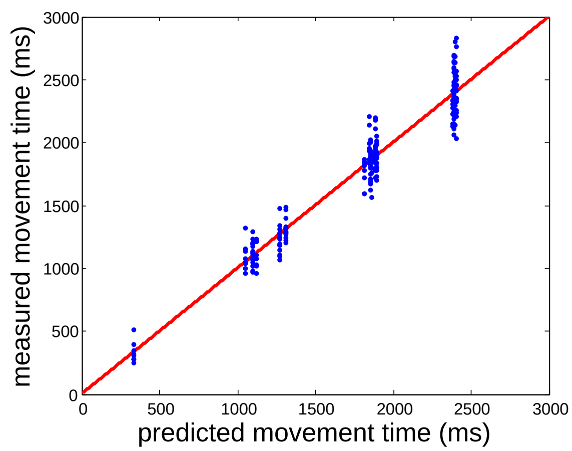


Figure 11: Model fitting with the empirical data of bimanual manipulations ($R^2 = 0.94, R = 0.97$).

However, we wondered how the model would perform with bimanual object manipulations, since the derivation of our model does not indicate any restrictions on manipulation types. Thus we repeated the conditions of the first experiment with another 10 participants but for bimanual manipulations, where participants were told to use their two index-fingers to perform the exact same task. Within all participants, 7 were male, 3 were females, and 3 were novice users. Their average measures included: age 22.4 ($SD = 1.5$), height 172 cm ($SD = 8.3$), hand length 17.8 cm ($SD = 1.5$), hand breath 10.6 cm ($SD = 0.9$), hand span 18.3 cm ($SD = 2.1$), and forearm length 43.8 cm ($SD = 2.4$).

We conducted a similar regression analysis with previous experiments for the empirical data collected in the above bimanual manipulation tasks. As Figure 11 shows, the results indicated that our model are still valid ($R^2 = 0.94$), where the estimated coefficients were: $a = 236$ ms, $b = 447$ ms/bit, $\alpha = 3.12$ pixel^2/degree^2, and $\beta = 1.02 \times 10^4$ pixel^2. Therefore, we believe that our model is very likely to be applicable for bimanual gestures, although not tested with larger range of conditions as the second experiment did.

10. Conclusion and Future Work

In this paper, a novel mathematical model for dual-finger multi-touch manipulation of 2D objects on touch-sensitive displays was proposed. Additionally two experiments were described that empirically validated the model. This model was derived by applying the weighted Euclidean distance to the index of difficulty of Fitts' law, which allows the incorporation of additional movement parameters pertaining to the multi-touch interaction, including translation, rotation and scaling. The resulting equation is a relation similar to Fitts' law, but that models a richer set of interactions. Regression analysis indicated that our model fits the empirical data very well (with R^2 and R values above 0.9). The model not only can predict the completion time of a multi-touch task, it can also be used to quantify the effort necessary to complete that task, which should make it possible to empirically compare devices and gestures in the same way that Fitts' law does. We also explored several aspects related to the multi-touch manipulation task and the proposed model, including the simultaneity of the movements and comparisons with other alternative models. However, we have not studied the model with empirical data of multiple levels of task tolerances and one shortcoming of this model is that the tolerance parameter W needs to be defined properly for a successful docking task. There are both pros and cons for the proposed accuracy measurements.

We can envision several opportunities for future work. We believe that this model can be used to improve multi-touch interfaces by supporting the empirical evaluation of gestures and gesture-based interfaces. This model could be extended to more elaborate physical activities, for example, higher degree of freedom interfaces employing an arbitrary number of fingers, or whole hand gestures. It would also be interesting to explore whether a model of higher dimensional path following may be created, in similar way to how Fitts' law was the basis from which Accot and Zhai (1997) modeled steering. We wish to conduct more experiments with a wider range of task tolerances to further validated our model, as only limited factors have been explored in this study. More practical further work is necessary to evaluate our model with multiple task tolerance levels, different devices, and various display sizes, in order to see what factors may effect the movement time and accuracy. Also, it is our intention to explore whether our model can be applied to interaction techniques that impose constraints in terms of allowing only a subset of the degrees of freedom, such as Rock&Rails (Wigdor et al., 2011).

Acknowledgments

We would like to express our thanks to those who participated in the experiments and to the reviewers for their time and helpful comments.

References

Accot, J., Zhai, S., 1997. Beyond fitts' law: models for trajectory-based HCI tasks. In: Proceedings of the ACM conference on Human factors in computing systems. pp. 295–302.

Accot, J., Zhai, S., 2003. Refining fitts' law models for bivariate pointing. In: Proceedings of the ACM conference on Human factors in computing systems. pp. 193–200.

Andersen, T. H., 2005. A simple movement time model for scrolling. In: CHI '05 extended abstracts. pp. 1180–1183.

Banovic, N., Li, F. C. Y., Dearman, D., Yatani, K., Truong, K. N., 2011. Design of unimanual multi-finger pie menu interaction. In: Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces. ITS '11. pp. 120–129.

Bi, X., Grossman, T., Matejka, J., Fitzmaurice, G., 2011. Magic desk: bringing multi-touch surfaces into desktop work. In: Proceedings of the ACM conference on Human factors in computing systems. pp. 2511–2520.

Bi, X., Li, Y., Zhai, S., 2013. Ffitts law: modeling finger touch with fitts' law. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '13. pp. 1363–1372.

Cao, X., Li, J. J., Balakrishnan, R., 2008a. Peephole pointing: modeling acquisition of dynamically revealed targets. In: Proceedings of the ACM conference on Human factors in computing systems. pp. 1699–1708.

Cao, X., Wilson, A., Balakrishnan, R., Hinckley, K., Hudson, S., 2008b. Shapetouch: Leveraging contact shape on interactive surfaces. In: IEEE International Workshop on Tabletop. pp. 129–136.

Casiez, G., Vogel, D., Balakrishnan, R., Cockburn, A., 2008. The impact of control-display gain on user performance in pointing tasks. Human-Computer Interaction 23 (3), 215–250.

Crossman, E. R., Goodeve, P. J., 1983. Feedback control of hand-movement and fitts' law. Quarterly Journal of Experimental Psychology. A 35 (Pt 2), 251–278.

Fitts, P. M., 1954. The information capacity of the human motor system in controlling the amplitude of movement. Journal of Experimental Psychology 47 (6), 381–391.

Forlines, C., Wigdor, D., Shen, C., Balakrishnan, R., 2007. Direct-touch vs. mouse input for tabletop displays. In: Proceedings of the ACM conference on Human factors in computing systems. pp. 647–656.

Freeman, D., Benko, H., Morris, M. R., Wigdor, D., 2009. Shadowguides: visualizations for in-situ learning of multi-touch and whole-hand gestures. In: Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces. pp. 165–172.

Grossman, T., Balakrishnan, R., 2004. Pointing at trivariate targets in 3d environments. In: Proceedings of the ACM conference on Human factors in computing systems. pp. 447–454.

Grossman, T., Balakrishnan, R., 2005. A probabilistic approach to modeling two-dimensional pointing. ACM Transactions on Computer-Human Interaction 12 (3), 435–459.

Guiard, Y., Beaudouin-lafon, M., 2004. Target acquisition in multiscale electronic worlds. International Journal of Human-Computer Studies 61 (6), 875–905.

Hinckley, K., Cutrell, E., Bathiche, S., Muss, T., 2002. Quantitative analysis of scrolling techniques. In: Proceedings of the ACM conference on Human factors in computing systems. pp. 65–72.

Hoggan, E., Nacenta, M., Kristensson, P. O., Williamson, J., Oulasvirta, A., Lehtiö, A., 2013a. Multi-touch pinch gestures: Performance and ergonomics. In: Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces. ITS '13. pp. 219–222.

Hoggan, E., Williamson, J., Oulasvirta, A., Nacenta, M., Kristensson, P. O., Lehtiö, A., 2013b. Multi-touch rotation gestures: Performance and ergonomics. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '13. pp. 3047–3050.

Knight, A. A., Dagnall, P. R., 1967. Precision in movements. Ergonomics 10 (3), 321–330.

Kruger, R., Carpendale, S., Scott, S. D., Tang, A., 2005. Fluid integration of rotation and translation. In: Proceedings of the ACM conference on Human factors in computing systems. pp. 601–610.

MacKenzie, C. L., Marteniuk, R. G., C. Dugas, D. L., Eickmeier, B., 1987. Three-dimensional movement trajectories in fitts' task: Implications for control. Quarterly Journal of Experimental Psychology. A 39 (4), 629–647.

MacKenzie, I. S., 1992. Fitts' law as a research and design tool in human-computer interaction. Human-Computer Interaction 7 (1), 91–139.

MacKenzie, I. S., Buxton, W., 1992. Extending fitts' law to two-dimensional tasks. In: Proceedings of the ACM conference on Human factors in computing systems. pp. 219–226.

MacKenzie, S. I., Soukoreff, W. R., 2002. Text entry for mobile computing: Models and methods, theory and practice. Human-Computer Interaction 17 (2), 147–198.

Masliah, M. R., Milgram, P., 2000. Measuring the allocation of control in a 6 degree-of-freedom docking experiment. In: Proceedings of the ACM conference on Human factors in computing systems. pp. 25–32.

Mason, A. H., Bryden, P. J., 2007. Coordination and concurrency in bimanual rotation tasks when moving away from and toward the body. Experimental Brain Research 183 (4), 541–556.

McGuffin, M., Balakrishnan, R., 2002. Acquisition of expanding targets. In: Proceedings of the ACM conference on Human factors in computing systems. pp. 57–64.

Nacenta, M. A., Baudisch, P., Benko, H., Wilson, A., 2009. Separability of spatial manipulations in multi-touch interfaces. In: Proceedings of Graphics Interface. pp. 175–182.

Nguyen, Q., Kipp, M., 2014. Orientation matters: Efficiency of translation-rotation multitouch tasks. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '14. pp. 2013–2016.

North, C., Dwyer, T., Lee, B., Fisher, D., Isenberg, P., Robertson, G., Inkpen, K., 2009. Understanding multi-touch manipulation for surface computing. In: Proceedings of the 12th IFIP TC 13 International Conference on Human-Computer Interaction. pp. 236–249.

Olafsdottir, H. B., Tsandilas, T., Appert, C., 2014. Prospective motor control on tabletops: Planning grasp for multitouch interaction. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '14. pp. 2893–2902.

Oulasvirta, A., 2014. Automated nonlinear regression modeling for hci. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '14. pp. 3899–3902.

Soukoreff, R. W., MacKenzie, I. S., 1995. Theoretical upper and lower bounds on typing speed using a stylus and soft keyboard. Behaviour and Information Technology 14 (6), 370–379.

Soukoreff, R. W., MacKenzie, I. S., 2004. Towards a standard for pointing device evaluation, perspectives on 27 years of fitts' law research in hci. International Journal of Human-Computer Studies 61 (6), 751–789.

Stoelen, M. F., Akin, D. L., 2010. Assessment of fitts' law for quantifying combined rotational and translational movements. Journal of Human Factors and Ergonomics Society 52 (1), 63–77.

Veit, M., Capobianco, A., Bechmann, D., 2009. Influence of degrees of freedom's manipulation on performances during orientation tasks in virtual reality environments. In: Proceedings of the ACM Symposium on Virtual Reality Software and Technology. pp. 51–58.

Wang, Y., MacKenzie, C. L., Summers, V. A., Booth, K. S., 1998. The structure of object transportation and orientation in human-computer interaction. In: Proceedings of the ACM conference on Human factors in computing systems. pp. 312–319.

Wigdor, D., Benko, H., Pella, J., Lombardo, J., Williams, S., 2011. Rock & rails: extending multi-touch interactions with shape gestures to enable precise spatial manipulations. In: Proceedings of the ACM conference on Human factors in computing systems. pp. 1581–1590.

Zhai, S., Milgram, P., 1998. Quantifying coordination in multiple dof movement and its application to evaluating 6 dof input devices. In: Proceedings of the ACM conference on Human factors in computing systems. pp. 320–327.

Zhao, J., Soukoreff, R. W., Balakrishnan, R., 2011. A model of multi-touch manipulation. In: Proceedings of the 2nd annual meeting of the Graphics, Animation, and New Media. GRAND.