Int. J. Human-Computer Studies 78 (2015) 68–80

Contents lists available at ScienceDirect

Int. J. Human-Computer Studies

journal homepage: www.elsevier.com/locate/ijhcs

# Exploring and modeling unimanual object manipulation on multi-touch displays ☆

Jian Zhao *, R. William Soukoreff, Ravin Balakrishnan

*Department of Computer Science, University of Toronto, 10 King's College Road, Toronto, ON, Canada, M5S 3G4*

ARTICLE INFO

*Article history:*
Received 14 May 2014
Received in revised form
21 February 2015
Accepted 23 February 2015
Communicated by J. LaViola
Available online 2 March 2015

*Keywords:*
Modeling
Human factors
Multi-touch displays
Virtual object manipulation
Fitts' law

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.



## 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) describe results based on limited preliminary data from our initial exploration of this problem. Specifically, our objective is to construct a

☆This paper has been recommended for acceptance by J. LaViola.

* Corresponding author. Tel.: +1 647 818 1919.
*E-mail addresses:* jianzhao@dgp.toronto.edu (J. Zhao), will@dgp.toronto.edu (R. William Soukoreff), ravin@dgp.toronto.edu (R. Balakrishnan).

http://dx.doi.org/10.1016/j.ijhcs.2015.02.011
1071-5819/

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 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 $W$,

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

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 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,

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

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