2012

Doctorate Thesis

Designing Touch-based Gesture Interactions

1138004　Huawei Tu

Advisor　Xiangshi Ren

September 2012

Course of Information Systems Engineering

Graduate School of Engineering, Kochi University of Technology

Abstract

Due to the powerful, efficient and convenient input properties, touch-based gestures have been widely employed to support a variety of interactive tasks. Touch-based gestures include pen gesture and finger gesture, which are movement trajectories of the contact point of the user's pen or finger on the touch sensitive surface. By means of touch-based gestures, users can select small targets, acquire remote targets, select menus, and entry text. Therefore, gesture-based interactions have been attracting widespread research interest in topics such as models for the prediction of gesture production time, algorithms for gesture recognition, and feedback for drawing gestures.

This thesis pays attention to two important issues about touch-based gesture interaction. One is regarding how to design gestures for touch-based interaction, with regard to different input forms (pen vs. finger), users of different ages (older users vs. younger users), and different entry sizes. The other is about gesture-based tasks, which concentrates on how to employ touch-based gestures in interactive activities.

For gesture-oriented design, three studies were conducted to improve touch-based gesture design. First, a study was conducted to quantify the differences and similarities between finger and pen gestures. The work proposed a methodology to investigate and quantify the performance of finger and pen gestures, and provided a solid foundation to apply principles, methods and findings from pen-based gesture design to finger-based gesture design. Second, a user-defined gesture study was conducted to compare user-defined gestures between younger people and older people in the context of pen input and finger input. The study aimed to understand the preferred gestures of both younger and older adults in the context of pen input and finger input. Third, as gesture entry size is an important factor for determining users' performance of gesture input, this thesis quantitatively investigated optimal finger-based entry size in touch-based mobile phones for two commonly used Chinese handwriting input styles: two-handed entry with the non-dominant hand holding the device and the index finger of the dominant hand entering characters; and one-handed entry with the dominant hand holding the device and the thumb of the dominant hand being used for character entry.

The experimental results and methodology can be employed in user interface design for gesture-based interaction in touch-based mobile phones.

To better understand the issue of gesture-based tasks, this thesis examined gesture performance in a document scrolling task in touch-based mobile phones and investigated the use of touch gestures to better support multi-user collaborative tasks on large tabletops as well. In the first study, this thesis quantitatively analyzed the performance of two scrolling techniques (flick and ring) for document navigation in touch-based mobile phones by means of three input methods (index finger, pen and thumb) in the context of sitting and walking postures. The work offered several insights for scrolling technique design for document navigation in touch-based mobile phones. The second study proposed Window Avatar, a window-based technique which allows the user to create a personal territory by means of hand shape gestures in multi-touch tabletop displays. Based on Window Avatar, this thesis presented a set of interaction techniques using shape gestures in combination with direct manipulations, so as to enhance user interaction on manipulation and collaboration.

In summary, this dissertation contributes to the field of gesture-based interaction in view of gesture-oriented design and gesture-based tasks. The conclusions drawn in this thesis and methodologies proposed in this thesis will be beneficial to future studies which aim to better explore touch-based gesture interactions, and to improve the design of touch-based gesture interactions.

Keywords

Touch-based interaction, gesture-based interaction, pen gesture, finger gesture, entry size, gesture-based task, gesture performance

Contents

List of Figures

List of Table

Chapter 1 General Introduction

1.1 Research Background

(a) (b) (c)

Figure 1. 1 Touch-based gesture interactions. (a) Flick gesture. (b) Zoom gesture. (c) Tap gesture.

Source: Google image.

Interaction (HCI) field. UI is responsible for the space where interaction between users and computers occurs. The goal of interaction between the user and the computer at UI is effective operation and control of the computer, and feedback from the computer which aids the user in making operational decisions. In general, there are three generations in the development of user interface techniques: command line interface (CLI), graphical user interface (GUI) and natural user interface (NUI) [70]. The popularity of Mac OS X, Microsoft Windows, and the X Window System has made GUI the most widely used interface in public life today. GUI allows the user to interact with the computer using metaphors (e.g., pictures and symbols), rather than having to memorize many complicated commands and to type them precisely, as with a command-line interface such as DOS. However, many drawbacks of GUI still seriously limit users' manipulation on computer. Users have to make frequent menu selections, button operations and commands input by means of the keyboard and the mouse, which will always result in a discrete operation procedure. Even though a task can be implemented in a single-pass, current software always breaks the implementation into a sequence of steps. NUI, as the next generation user interface, has generated considerable recent research interest. NUI refers to a user interface which is (1) effectively invisible, or becomes invisible with successive learned interactions, to its users, and (2) is based on nature or natural elements (i.e. physics, also known as Natural Philosophy) [70].

As a powerful, efficient and convenient input style, gestures are one desirable feature of NUI. “A gesture is a motion of the body that contains information.” [51] Gestures in HCI can be classified into two categories: two-dimensional gestures *surface gestures* [105] (see Figure 1. 1) and three-dimensional gestures *motion gestures* [81]. For surface gestures, users can draw gestures on the touch screen in two dimensions. For example, flick gesture, which requires the user to contact the digitizer in a quick flicking motion, and zoom gesture, which is used to enlarge or reduce a object by bringing fingers in the opposite direction or closer together. Motion gestures can enable users to interact with a device, in three dimensions, by translating or rotating the device, or by moving the hands, face or other parts of the body without holding a device.

The advantages of gesture-based interactions are two folds: (1) the user can quickly articulate the gesture for a command or a word by recalling from memory, rather than by selecting an icon with looking much at the icon, which may be a time-consuming process; (2) compared to tapping on the icon, gesture input is a form of more fluid movement closer to drawing, hence introducing a natural input style for the user. The above attributes of gestures can meet the requirements of NUI and can contribute to the growth in popularity of NUI. Therefore, gestures have been attracting widespread interest in HCI research field.

This thesis focuses specifically on surface gesture. Due to the rapid growth of touch screen devices, gestures on touch screens are an increasingly important interaction modality. Touch-based gesture is a movement trajectory of the contact point of the user’s pen or finger on the touch sensitive surface; finger gesture and pen gesture are two main forms of touch gesture. Gestures can be used in a number of application scenarios for touch screens. A typical application is the use of gestures in target selection task, which is regarded as a fundamental task in HCI. Lü and Li [61] showed how to apply gestures to select small targets in user interfaces. Gestures can also be employed to acquire remote targets in large interactive displays [10]. In addition, users can employ gestures to perform command selection as an alternative to pull-down menus in pen-based interfaces [5], [45]. In addition to the above applications, one noteworthy application of gestures is text entry.

Shapewriter is such a technique, which is a novel form of writing that uses pen strokes on graphical keyboards to write text, can enable users to enter text efficiently at a faster rate than previously possible on mobile phones, handheld computers and other mobile devices [46], [115].

As a result, in view of a wide variety of applications with gestures, gestures have attracted widespread research interest in topics of recognition algorithm, feedback and models. We will review the related work in Chapter Two.

1.2 Objectives and Research Issues

This thesis pays attention to two important questions about touch-based gesture interaction. One is how to design gestures for touch-based interaction, with regard to different input forms (pen vs. finger), users of different ages (older users vs. younger users), and different entry sizes. The other is how to employ gestures in interactive activities.

For gesture design, this thesis explores three aspects in touch-based gesture interaction: input forms, users of different ages and entry sizes. (1) Regarding input forms, this thesis examined the performance of pen gesture and finger gesture. Previous studies have investigated gesture performance with regard to pen input and finger input respectively. As the stylus (pen) has been the primary implement for drawing stroke gestures on touch screens, such as that of PDAs, past stroke gesture research has been focused on the digital pen as the drawing implement. And most stroke gesture HCI research work published to date, such as [4], [5], [6], [14], [26], [45], [60], [79], [115] has been based on data collected from gestures produced with high quality inductive digital styli. However, recent commercial product design has tended to avoid the use of the pen with a view to user convenience and simplicity. Hence, a major current focus in gesture design refers to finger gesture design [1], [56], [61], [71], [105], as well as the combinational use of finger and pen gestures [28], [116]. Nevertheless, little research has paid attention to the analysis of pen gesture and finger gesture together, with the consideration that the two kinds of gestures have some common and distinct features. Treating finger gesture and pen gesture as two independent parts of touch gesture can not deeply explore their characteristics for the designer, which would hinder efficient design and evaluation of touch gesture based interaction. Pen gesture and finger gesture have some obvious differences and similarities as follows: first, pen input and finger input are all isotonic (zero or constant resistance) and position control input styles; second, pen input is more precious than finger input, but finger input is more direct than pen input. However, it remains unclear of the deep-seated differences and similarities between finger gesture and pen gesture, e.g., in precision, size, and other gesture characteristics. Understanding the similarities and differences plays an important role in touch-based gesture design; finger gesture design and pen gesture design can learn from each other. This thesis aims to fill in the above gap of gesture research, with a focus on how quantitatively similar or different finger gestures are from pen gestures, as well as on how to apply principles, methods and findings from pen-based gesture design to finger-based gesture design. (2) In respect to users of different ages for gesture input, it is well known that perceptual, cognitive and motor deficits result in many older adults experiencing greater difficulties performing computer-related tasks than younger adults [32], [34], [63]. However, because younger adults are the main consumer groups of interactive devices, current gesture design aims to meet their needs, while ignores the needs of older people. Most gesture-based interfaces provided few or no accessibility features for older people, leaving the interfaces largely unusable for that age group. Although this is a serious problem, there has been very little work on the investigation of gesture performance involving the consideration of age related factors. It is more important to design appropriate gestures for older adults than for younger adults in two reasons. First, due to the perceptual, cognitive and motor deficits of older adults, it would take longer time for them to learn how to use gestures. Hence, appropriate gestures can allow older people to learn and use more easily. Second, compared to younger adults, older adults are likely to feel deeper frustration when meeting the failure to perform gestures, which hinders older adults from continuing use gestures. Therefore, it is vital to understand how older people prefer to use gestures, and to design better gesture-based interactions for older people. (3) Entry sizes play an important role in users' performance of handwriting entry [79]. For interactive devices, such as mobile phones, which have a small screen, the screen area restricts the handwriting entry area size. In order to design a rational screen layout which can display more information and also allow users to draw gestures with ease and high efficiency, it is important to determine the optimal entry area for gesture input. There are other properties of touch-based gesture interaction which are impossible to explore all at once.

This thesis concentrates on the above three topics, since they are fundamental and their impact on performance is important to the design of touch-based gestures. Once we establish a basic understanding of these properties, further research issues will be exposed and can be pursued in future studies.

For gesture-based task, this thesis examined gesture performance in a document scrolling task in touch-based mobile phones and also designed a set of gestures to better support multi-user collaborative tasks on large tabletops. The document scrolling task was selected for analysis because this is a common task in HCI field [27]. Two scrolling techniques, flick and ring, which are two important scrolling gestures for document navigation, are analyzed here. Examining the advantages and disadvantages of these two scrolling gestures would be beneficial to scrolling technique design. The other interactive task examined in the study includes individual and collaborative tasks in multi-touch tabletop displays. Multi-touch tabletops have been widely employed in individual and collaborative tasks. However, they suffer from some drawbacks such as about the orientation of the tabletop and remote targets acquisition [10], [47], [86]. To improve interactions in multi-touch tabletop displays, a set of interaction techniques were presented using shape gestures in combination with direct manipulations, so as to enhance user interaction on manipulation and collaboration. These two studies aimed to better employ gesture-based interactions in HCI tasks.

1.3 Dissertation Overview

The purpose of this dissertation is to examine the performance of touch-based gesture interactions, so as to better design touch gestures. The structure of this dissertation is shown in Table 1. 1. As a first step, an introduction is presented to describe the research background and the motivation of this dissertation (Chapter 1). After the introduction, a detailed review of previous studies about touch-based gesture interactions is presented (Chapter 2). Then, five gesture-related studies which aim to improve touch-based gesture interaction are introduced in five chapters in the context of gesture-oriented design and gesture-based task. Among these five studies, three of them are referred to gesture-oriented design, which will be described in Chapter 3, 4 and 5 respectively.

First, Chapter 3 introduces a study which was carried out to compare the performance of finger stroke gesture and pen stroke gesture in terms of a set of gesture features. Second, a study (Chapter 4) was conducted to examine the performance of pen gesture and finger gesture for older people and younger people by means of a user-defined approach. In addition, Chapter 5 shows a study regarding the optimal entry box size for Chinese characters (as complex gestures) in touch-based mobile phones. For gesture-based task, a study was conducted to examine the performance of two scrolling gestures with pen input and finger input (Chapter 6). In addition, a study regarding gesture-based interaction for multi-touch tabletops is presented (Chapter 7). Each chapter begins with a research motivation, followed by a literature review. Experiments are then detailed with the description of carefully designed representative tasks. Conclusions are drawn based upon rigorous statistical analysis of the experimental results. Through these five studies, a number of approaches are presented to reveal the properties of gesture-based interactions, and also are showed on how to better apply touch-based gesture interactions to user interface design. After the five chapters, the conclusions of the dissertation are presented and future directions are also discussed (Chapter 8). Overall, the purpose of this thesis is to improve the performance of touch-based gesture interactions.

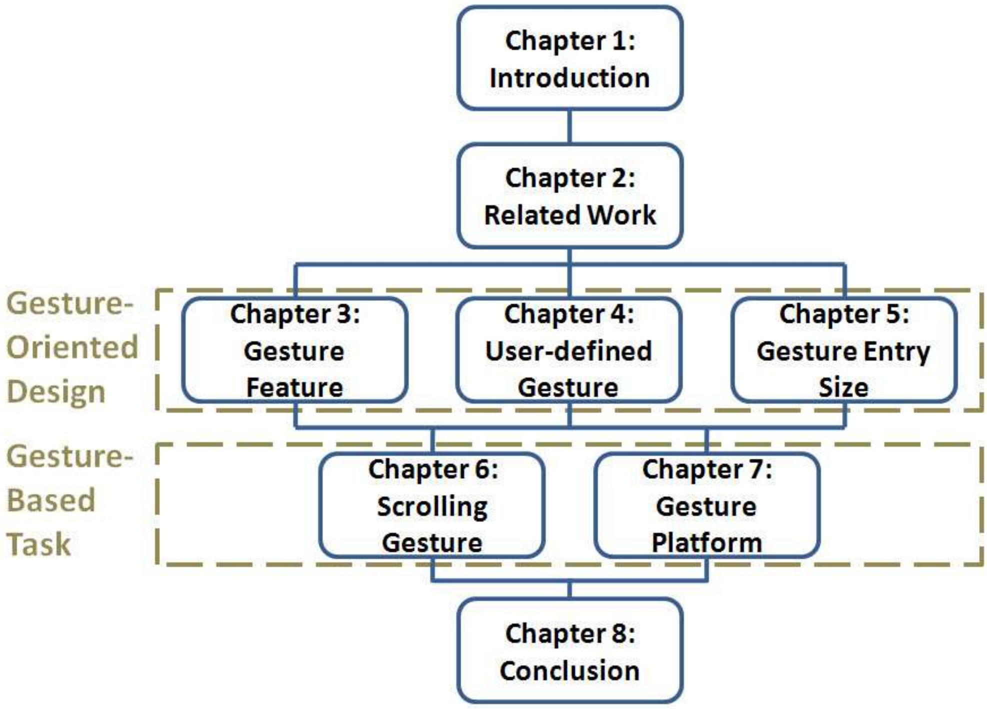


Table 1. 1 The dissertation structure.

Chapter 2 Literature Review

This chapter reviews past studies regarding touch-based gesture interactions from two aspects: (1) the history of touch-based gesture interactions; (2) current research topics of touch-based gesture interactions. In the first aspect, this thesis reviews the development of gesture-based interaction, with a specific focus on two important changes of touch-based gesture interactions in the important paradigm shift. In the second aspect, this thesis summarizes three important research issues in gesture-related research field: gesture recognition algorithm, motor control complexity of stroke gestures, and feedback in gesture interfaces. The review of previous studies demonstrates the significances of the studies presented in this thesis, and also provides several beneficial research methodologies to this thesis.

2.1 History of Touch-based Gesture Interactions

Touch-screen gesture research has attracted much attention for many years in HCI research. As early as in 1963, Sutherland [96] conducted an early project, Sketchpad, which is treated as one of the beginnings of human-computer interaction research. The project aimed to use stroke gestures in graphical human-machine communication. Newman and Sproull [68] prominently featured stroke gestures as an input mechanism and described in detail how to implement a rudimentary stroke gesture recognizer. Since then, surface gestures have attracted wide research interests in HCI research field [15], [73], [111]. These studies explored the advantages of gesture-based interaction and contributed to develop gesture-based interaction in user interface design.

It is beyond the scope of this thesis to conduct a detailed review of the history of touch-based gesture interactions. Instead, we would like to highlight two important changes of gesture based interactions in the history: (1) the shift from pen gestures to finger gestures; (2) the change from stroke gestures to multi-touch gestures. Through the description of these two changes, we aim to show the significance of the studies presented in this thesis.

As the stylus (pen) has been the primary implement for drawing stroke gestures on touch screens, such as that of PDAs, past stroke gesture research has been focused on the digital pen as the drawing implement. And most stroke gesture HCI research work published to date, such as [4], [5], [6], [14], [26], [45], [60], [79], [115] has been based on data collected from gestures produced with high quality inductive digital styli. However, recent commercial product design has tended to avoid the use of the pen with a view to user convenience and simplicity. Hence, a major current focus in gesture design refers to finger gesture design [1], [56], [61], [71], [105], as well as the combinational use of finger and pen gestures [28], [116]. This kind of shift raises several research questions regarding gesture-based interactions. This thesis deeply explored one of these research issues: the differences and similarities between pen and finger gestures, and discussed several research guidelines in the context of the shift regarding gesture entry style.

Past surface gesture studies focused on stroke gestures, which are usually drawn by a pen or a finger [1], [6], [9], [14], [60], [115]. Advances in touch screen allow users to draw touch gestures with multi fingers or combinational input styles with pens and fingers. Higher degree of freedom enables the user to easily perform some gestures, such as using zooming gesture to enlarge targets. Using multi fingers to perform such gestures is consistent with user experiences in real world situations. This input attribute meets the requirement of NUI. Therefore, multi touch gestures are becoming more and more widely used in commercial products. Also, the research of touch-based gesture interactions became interested in multi-touch gestures [11], [13], [24], [23], [37], [39], [40], [61], [82], [98], [103]. In this thesis, we examined the performance of stroke gestures and multi touch gestures respectively, which aims to improve the design of stroke gestures and multi touch gestures.

The above two changes indicate that gesture-based interactions, as a natural and convenient interaction style, are gaining popularity in interactive device design. As we have summarized previously, the advantages of gesture-based interactions are two folds: (1) the user can quickly articulate the gesture for a command or a word by recalling from memory; (2) gesture input is a form of more fluid movement closer to drawing, hence introducing a natural input style for the user. The above attributes of gestures can meet the requirement of NUI and can contribute to the growth in popularity of NUI.

2.2 Research Topics Related to Touch Gestures

This thesis conducted a detailed literature review of touch-based gesture in current HCI research and sorted them into three major topics: gesture recognition algorithm, motor control complexity of stroke gestures, and feedback in gesture interfaces. Through reviewing these three fundamental aspects of gesture research in HCI field, we aim to provide some valuable methodologies and guidelines for the studies described in the thesis.

2.2.1 Gesture Recognition Algorithm

A major research interest in gesture interface is the design and development of gesture recognition algorithms. Single stroke gesture recognition was a critical part of early handwriting recognition systems. The current practices methods of gesture recognition can be classified into two categories. One is training-based methods and the other is template matching methods.

Training-based methods include Hidden Markov Models (HMMs) [3], [12], [90], neural networks [75], feature-based statistical classifiers [17], [80], dynamic programming [66], [97] and ad-hoc heuristic recognizers [17], [69], [110]. These approaches represent a stroke gesture as an n-dimensional vector and use a training set to partition the n-dimensional space into multiple gesture classes. Among these approaches, one noteworthy approach is the recognizer proposed by Rubine [80], which encodes a gesture as a vector of 13 features and uses a covariance matrix to partition this 13-dimensional space. This recognizer has been used in many gesture research projects [30], [54], [59], [67].

As a simple and easy to implement approach, template matching based recognizer has attracted much research attention and has been widely employed in many gesture-related studies [5], [45], [57], [108], [115]. For gesture recognition, template-based methods compute the distance between user's gesture input and a list of gesture templates, after rotating, scaling and translating. This thesis used proportional shape distance as a feature in the first study and examined the performance of pen and finger gesture in terms of the feature.

2.2.2 Motor Control Complexity of Stroke Gestures

Some studies paid attention to the research on quantitative models of human performance in producing gestures that can characterize the efficiency of a given gesture or a gesture set. Such models may be useful in the design and evaluation of existing or future gesture interfaces by quantitatively predicting their efficiency before running extensive user studies. This thesis reviews three noteworthy studies as follows.

Isokoski [36] proposed a model for stroke gestures that used the number of approximating straight line segments in a gesture as a predictor of complexity correlating to production time. The underlying assumption is that drawing a straight line segment takes constant time, regardless of the length of the segment. The model’s best correlation result was $R^2 = 0.85$ on Unistroke characters [25], and it achieved R^2 between 0.5 and 0.8 for other gesture sets. However, it is difficult to accurately define the number of straight line segments needed to approximate a curved gesture. Furthermore, it does not provide an estimation of the magnitude of the actual production time.

Viviani and colleagues [101], [102] investigated human handwriting and drawing behavior in terms of instant movement velocity as a function of curvature, and proposed a power-formed model.

$$V = KR^{\beta} \tag{2.1}$$

, where V is the instant (tangential) velocity of movement; R is the radius of curvature C (R=1/C), and K and β are constants of the model.

CLC model, proposed by Cao and Zhai [14], is a quantitative human performance model of making single-stroke pen gestures within certain error constraints in terms of production time. In CLC model, a stroke gesture is decomposed into three classes of gesture elements: curves, line segments and corners. The total production time of a gesture is computed as the sum of all time durations of producing all gesture segments:

$$\mathrm{T}=\sum \mathrm{T(line)}+\sum \mathrm{T(coner)}+\sum \mathrm{T(curve)} \tag{2.2}$$

Results showed that high correlation with empirical data from a variety of gesture sets were achieved, with greater than 0.9 R^2 value in all cases. Thus, CLC model can serve as a foundation for the design and evaluation of existing and future gesture-based user interfaces at the basic motor control efficiency level. In the first study of this thesis, CLC model was used to predict gesture production time, so as to determine gesture complexity with other considerations of gesture length and gesture appearance.

2.2.3 Feedback in Gesture Interfaces

Feedback plays an important role in the performance of HCI techniques. Targeted this point, the importance of feedback in gesture based interaction has been deeply explored. Andersen and Zhai [4] investigated the use of auditory feedback in pen-gesture interfaces. They found that it was difficult to enhance the performance of gaining performance and learning advantage through auditory feedback but a few simple functions such as indicating the pen-gesture recognition results can be achieved. Visual feedback can serve as an effective method for learning gesture-based command sets [6]. EdgeWrite [107] is an example of exploiting haptic assistance in gesture production. It is a text entry method for handheld devices designed to provide stability of motion for people with motor impairments.

In summary, the review shows that touch-based gesture interactions are being promoted with the development of user interface. And this interaction style is gaining popularity in Natural User Interface (NUI). The review also shows some useful methodologies and findings for the studies described in the thesis.

Chapter 3 A Comparative Evaluation of Finger and Pen Stroke Gestures

3.1 Introduction

Due to the rapid growth of touch screen devices, stroke gestures on touch screens are an increasingly important interaction modality. Until recently, the stylus (pen) has been the primary implement for drawing stroke gestures on touch screens. However, today's preferred implement for tapping and gesturing on touch screens is the finger or fingers.

Past stroke gesture research has been focused on the digital pen as the drawing implement. Most stroke gesture HCI research work published to date, such as [5], [6], [46], [49], [50], [60], [108] has been based on data collected from gestures produced with high quality inductive digital styli. It is questionable whether and how well these results apply to finger drawn gestures. Our investigations looked at the differences and similarities between finger and pen stroke gestures both of which have been neglected in the literature. There is a clear need to identify and characterize these differences where they are present. For example, if we know finger gestures are particularly poor at producing certain types of features, then future research and product design should exploit such knowledge and avoid relying on these features in their recognition algorithm and gesture set design. Understanding the quantitative difference between finger strokes and pen strokes can provide a foundation for differential designs of pen and finger interface or combinational designs of pen and finger input in the same interface [28].

To our knowledge little has been done in the HCI research community to address these pressing questions. We see gesture interfaces such as gesture keyboards but they were initially designed with the pen in mind [115] and have been increasingly transformed to finger use [116]. However, the costs and benefits of this unevaluated adaption switch are not known beyond anecdotal subjective impressions. A scientific approach, such as the one presented in this study, has been pending for too long.

We set out to perform the first systematic comparative investigation between pen vs. finger gestures. We asked participants to draw a set of stroke gestures with a finger and a pen respectively as shown in Table 3. 1. We then processed and analyzed the drawn gestures according to a set of measures and features that are either most basic (such as stroke articulation time) or that are likely to differentiate finger gestures from pen gestures (such as the precision of corner production). We then drew a set of conclusions that characterize the differences and similarities between finger vs. pen gestures.

3.2 Related Work

There is a large body of HCI research on gesture interfaces, e.g. [10], [64], [105] for finger gesture and [5], [6], [46], [50], [60], [108] for pen gesture. It is unnecessary and beyond the scope of this study to review that literature here. Instead, we only highlight a few lines of work that bear direct relevance to the questions we addressed and the methods we used in addressing them.

3.2.1 Human Motor Control Theory

Historically, the study on how humans control their motor behavior has centered on the debate between the centrists and the peripheralists among motor control theorists [85]. The centrists tended to view motor control behavior as an inside-out process, driven by “motor programs” from human internal representations. In contrast, peripheralists tended to emphasize motor control behavior as regulated by outside-in feedback from the environment. To our current questions regarding finger vs. pen gesture differences, a centralist would suggest that there is little difference between finger and pen gestures since their production are both driven from internal representations, as is indeed proposed in the effector independence theory [104] concerning writing. A peripheralist however would argue that the different feel and interaction with the touch screen surface afforded by the pen vs. the bare finger would impact how a gesture is produced.

3.2.2 Gesture Models

The complexity of a stroke gesture may have an impact on the difference between finger and pen gestures. Conceivably fingers are good (enough) at producing simple gestures. How to measure and characterize gesture complexity is a research topic. Simple measures such as the length or the number of line segments [36] in a stroke gesture may serve as complexity indicators. A more formal model, the CLC model [14] that computes a gesture's production time based on sub models of curve, line, and corner production, is a more rigorous characterization of gesture complexity. We used the CLC model as a verification method in the design of our experiment.

3.2.3 Gesture Measurements and Features

Blagojevic et al. [8] categorized a feature library of ink gestures and used this library with attribute selection algorithms to choose good features for gesture recognition. Their work revealed that feature selection can significantly improve recognition rates, which demonstrated the importance of selecting good features for gesture recognition. However, their study did not pay attention to either finger stroke gestures or the difference and similarities between finger and pen gestures. Our study investigates the differences and similarities between finger and pen gestures, so as to find “finger friendly” features for finger gesture design and recognition.

Gesture recognition algorithms inevitably use a set of features to classify user input. These features can all be potentially used as measures of finger and pen gesture difference. For example, Andersen and Zhai [4] developed a set of measures to characterize gesture difference. SHARK[2] used proportional shape distance (*PSD*) as a key feature in classifying the user's input on a gesture keyboard [46]. The *PSD* feature was more generally studied in Wobbrock et al. [108] showing that it produces comparable or stronger results than the well known Rubine recognizer [80] that combines a set of features through data training. Long et al. [60] used a set of features mostly taken from the Rubine recognizer.

Five features, namely proportional shape distance, indicative angle difference, time, speed, distance between the first and last points, from the above cited papers [4], [46], [60], [80], [108] appeared to be most relevant to the research questions we wanted to address here.

3.3 Gesture Used in the Experiment

In order to identify differences between finger and pen gestures, we designed and selected a set of gesture prototypes. Our goal was to have a small gesture set that covers a wide range of gestures across different categories.

3.3.1 Gesture Categories

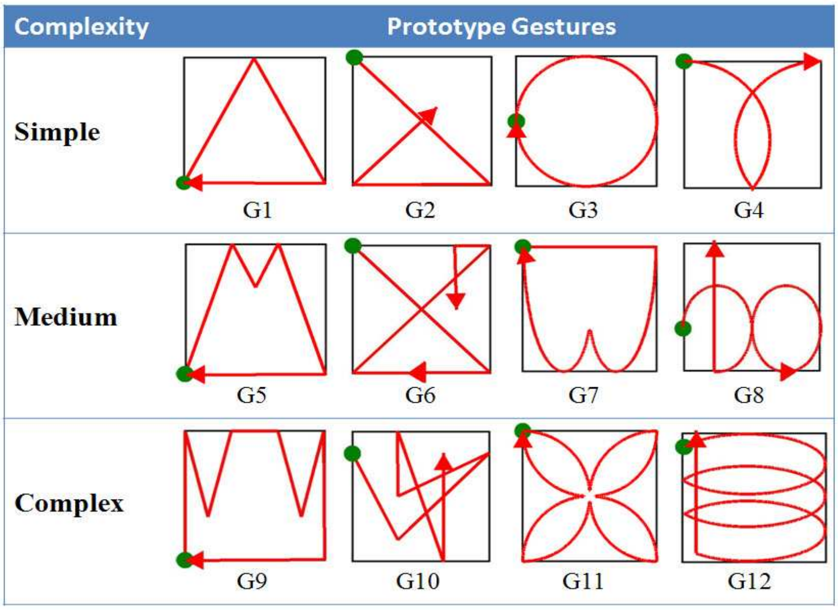


Table 3. 1 Prototype gesture categories. The green dot signifies the starting point of a gesture, and the arrow denotes the end point and the direction of a gesture.

Twelve gestures were used in our experiment. Their prototypes are shown in Table 3. 1. Five of them were selected from previous studies (G1 [4], G2 [80], G3 [4], G4 [5], [6], G10 [115]). Four were designed based on previous studies (G5 [80], G6 [4], [5]}, G7 [4], G12 [60]). G8, G9 and G11 were newly designed for this study.

Based on visual appearance in terms of the number of corners, curves and line segments, the gestures were divided into three groups according to their levels of complexity, i.e., Simple, Medium and Complex as shown in Table 3. 1. These classifications were also supported by simple complexity measures such as length and by their predicted production time. The length of a prototype gesture with the bounding box 3.0 × 3.0 cm was the sum of the distance between adjacent points in the prototype gesture's trajectory. The predicted production time for a prototype gesture with the bounding box 3.0 × 3.0 cm was calculated by the CLC model [14]. For example, the length and expected time for G1 are 9.7 cm and 1006 ms respectively, while for G12, the length and expected time are 22.67 cm and 2829 ms respectively.

These gestures also vary in characteristics. Gestures G1, G2, G5, G6, G9 and G10 were composed of corners and straight lines, and Gestures G3, G4, G7, G8, G11 and G12 were mainly composed of corners and curves. Gestures G1, G3, G5, G7, G9 and G11 are closed gestures because their prototypes start and end in the same position. The rest of the gestures in the set are open gestures.

Gestures G2, G4, G6, G8, G10 and G12 contain intersections, and the other gestures do not. The number of interaction points generally increases with gesture complexity. Gestures G2 and G4 have one interaction point each, G6 and G8 have two interaction points each, G10 has four, and G12 has seven interaction points. Gestures G1, G3, G4, G5, G7, G9 and G11 are symmetrical about the Y axis. The others are asymmetrical.

3.3.2 Target Gesture Size

Intuitively, stroke gestures can be more easily produced in smaller sizes with the pen than with the finger. This led us to repeat the same set of gestures in three different target sizes and ask the participants to reproduce them accordingly. The target gesture size of a prototype gesture was defined as the area in cm^2 of the target gesture's bounding box. From past research we know that pen gestures can be produced in rather small sizes.

According to Ren and Zhou [79], the bounding box size 1.5 × 1.5 cm in length was set up as a baseline in our experiment, which should be rather comfortable for pen gesturing and we suspected that it would be more challenging for finger gesturing. To evaluate the gesture size factor, we also set up the medium (3.0 × 3.0 cm) and large target gesture sizes (4.5 × 4.5 cm) respectively. We expected these two sizes would be less challenging for finger gesturing.

3.4 Experiment

3.4.1 Participants

Fifteen volunteers, twelve males and three females, from 20 to 30 years of age, participated in the experiment. All participants were right-handed. Ten of them had prior experience using stylus, and also with finger operation. Three of them had prior experience with finger operation only on touch screen devices. The other two participants had no prior experience operating digital screens with either stylus or finger.

3.4.2 Apparatus

The study was conducted on a HP touchsmart tx2 tablet computer. The screen size was 12.1 inches and its resolution was 1280 × 800 pixels, which means the pixel pitch was 0.204 mm. The most important reason we chose this computer as the experimental apparatus was that it has two touch sensing mechanisms (one capacitive and the other inductive), hence supporting both pen and finger gestures [33]. This ensured that we measured finger and pen gestures under the same set of conditions and form factors. During the experiment the computer was folded in tablet mode and laid on the table with the screen facing upward.

3.4.3 Task and Procedure

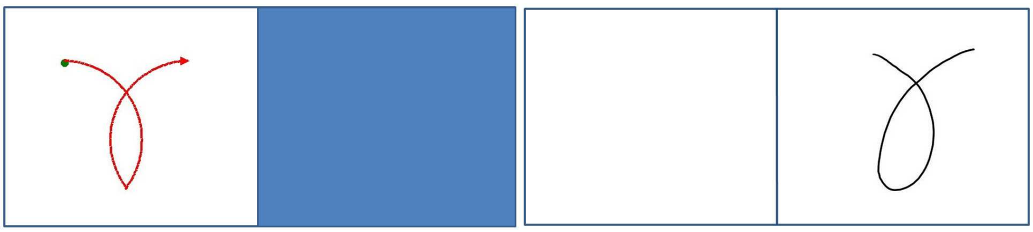


Figure 3. 1 The display for gesture input.

The goal of the experimental task design was to simulate how people draw gestures from their memory instead of copying or tracing a template.

Similar to the experimental design of [4], participants were asked to draw the gesture from memory as accurately as possible at a normal writing speed, using the pen and index finger of the dominant hand, after being shown the target gesture. As shown in Figure 3. 1, the experiment window was divided into a display area and a gesture input area. In each test trial, a gesture prototype was displayed in the left window for 1.5 seconds, with a dot and an arrowhead indicating the starting point and ending point respectively; meanwhile the right window was hidden by blue color (see Figure 3. 1 left).

After 1.5 seconds, the gesture prototype disappeared, along with the blue color in the gesturing area, prompting the participant to draw the same gesture in the right window (see Figure 3. 1 right). Pilot studies indicated that after a training period, this time period is long enough to allow participants to remember both the size and overall shape of the target gesture.

The experiment consisted of a training phase and an experimental phase. In the training phase, participants were first taught how to perform the experimental task. Then they were asked to draw the twelve gestures in three sizes using the pen and finger respectively as practice. In this training phase, the gesture prototype remained in the left window till the entire trial was completed. In the test phase of this within-subject experiment, each participant completed four blocks of all gestures in three sizes in two drawing implement conditions: pen vs. finger. Within each block, the order of the twelve gestures in three different sizes was randomized. In summary, the experiment data collection consisted of:

15 subjects ×

2 implements ×

4 blocks ×

12 gestures ×

3 target gesture sizes ×

= 4320 drawing trials

At the end of the experiment, a questionnaire was administrated to gather subjective opinions.

Participants were asked to rate *pen input* and *finger input* on 7-point Likert Scales regarding *speed, accuracy* and *hand fatigue* (7 for highest preference, and 1 for lowest preference).

We defined a set of dependent variables, including time, accuracy and shape similarities between user drawn gestures and the prototypes. For ease of comprehension and brevity, we deferred the formal definition of each dependent variable to the next section and placed it immediately before the corresponding experimental results.

3.5 Feature Selection

As described in the section “Related Work”, a lot of stroke features have been studied. For the purpose of our study, we chose five features from the literature and designed four new features (see Table 3. 2). We suspected that all these features may reveal differences between finger and pen gestures. With features F1 and F2, the pen or the finger used as the drawing implement may lead to different performance due to either friction or dexterity differences. In addition, because the pen tip is sharper and allows more precision than the fingertip, the finger may result in different performance with respect to F3, F4, F5, F6, F7, F8 and F9.

Inspired by the feature classification method in [80], each feature was classified manually along two dimensions: algebraical property feature and geometric shape feature (see Table 3. 2). As a basic measure, the algebraical property feature represents the basic features of a gesture, including stroke time, movement speed and size ratio. The geometric shape feature consists of the local shape feature and the global shape geometry feature. It focuses on what a gesture looks like.

Feature Categories	Measures	Features
Algebraical Property Feature	Basic Measure	F1. Time Performance F2. Average Speed F3. Size Ratio
Geometric Shape Feature	Local Shape Measure	F4. Aperture between the Start Point and the End Point of Closed Gestures F5. Indicative Angle Difference between Drawn Gesture and Target Gesture F6. Corner Shape Distance
	Global Shape Measure	F7. Axial Symmetry F8. Proportional Shape Distance F9. Intersecting Points Deviation

Table 3. 2 Feature categories.

We also conducted a pilot study to find differences between finger gestures and pen gestures by means of a set of commonly used gestures.

We chose seven gestures from *Graffiti*, which is a single-stroke shorthand handwriting set widely employed in PDAs. The seven gestures denote the character “a”, “b”, “c”, “e”, “d”, “j” and “%” respectively. In addition, we selected a square gesture and a five-pointed star gesture.

Six participants took part in the pilot study. The experiment procedure was similar to that introduced in the section “Task and Procedure”.

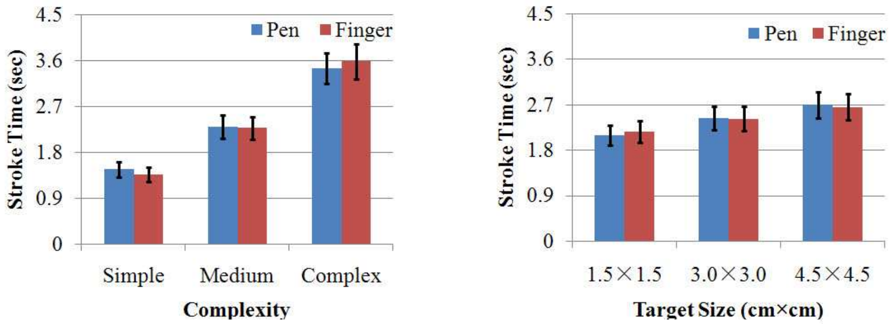
The experimental data were assessed in terms of the 63 features for single stroke gestures used in [8]. We found that the features with differences between finger and pen gestures mainly referred to movement speed, size and curvature, which had already been included in Table 3. 2. Hence, it can be regarded that this feature set shown in Table 3. 2 can demonstrate the main differences between finger gestures and pen gestures.

3.6 Results and Analysis

In this section, we discuss the experimental results in terms of the gesture features listed in Table 3. 2. Recall that each participant performed four blocks of trials in the experiment, we first checked the learning effect on stroke articulation time over the four blocks of trials to see if the data we collected had reached a level of stability. As to results, the participants' performance began to stabilize in the second block of trials for finger strokes and in the third block of trials for pen strokes. Therefore, data in the third and fourth blocks were applied to the rest of our analysis for pen strokes, and data in the second, third and fourth blocks were applied to the rest of our analysis for finger strokes.

3.6.1 Basic Measures

Time Performance



(a) (b)

Figure 3. 2 Stroke articulation time for each implement in different (a) complexities and (b) target gesture sizes. Error bars represent 0.95 confidence interval.

Stroke articulation time was defined as the duration from the moment the pen or finger touched the screen to the moment the pen or the finger was lifted from the screen. This is a basic measure of stroke performance. Conceivably, there could be a difference in this measure between the pen and the finger as the drawing implement due to either friction or dexterity differences. However, repeated measures ANOVA showed that the drawing implement (pen vs. finger) had no significant main effect on stroke articulation time. The mean stroke articulation time was 2408 ms in the pen condition and 2414 ms in the finger condition.

Other independent variables influenced stroke articulation time. As expected, the level of gesture complexity had a significant main effect on mean stroke articulation time ($F_{2, 28}$ = 127.88, p < 0.001). The target gesture size also had a significant main effect on mean stroke articulation time ($F_{2, 28}$ =

67.14, $p < 0.001$). There was a strong interaction between implement and complexity ($F_{2, 28} = 8.44$, $p < 0.01$). As shown in Figure 3. 2a, the mean stroke articulation time of the pen was longer than that of the finger in drawing simple gestures (1468 ms vs. 1370 ms), slightly longer in drawing gestures of medium level complexity (2306 ms vs. 2284 ms), but shorter in drawing complex gestures (3451 ms vs. 3587 ms). Also, there was a significant interaction between implement and target gesture size ($F_{2, 28} = 12.08$, $p < 0.001$). Figure 3. 2b illustrates that for small target size, the mean stroke articulation time achieved with pen input (2092 ms) was shorter than that for finger input (2170 ms). However, for medium and large target sizes, pen input led to longer stroke articulation time than finger input (2435 ms vs. 2420 ms for medium size, and 2698 ms vs. 2650 ms for large size). The pen tended to be slightly slower in drawing simple gestures and large size gestures.

Average Speed

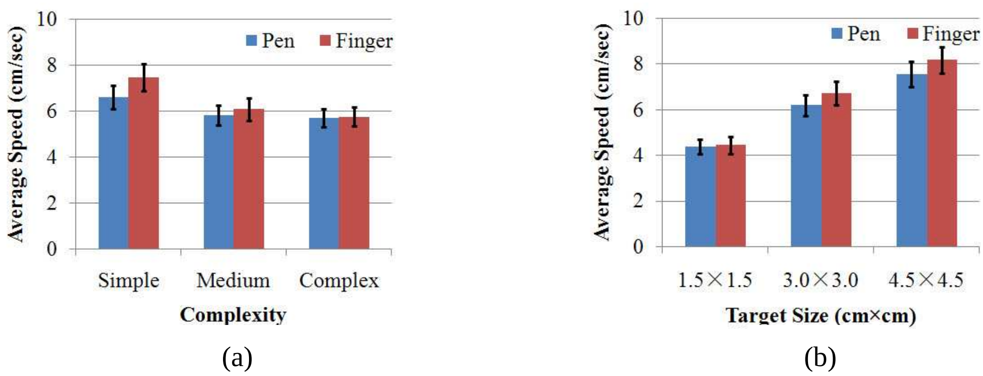


Figure 3. 3 Average speed for each implement in different (a) complexities and (b) target gesture sizes

The average speed, calculated by the ratio of the gesture length and the stroke articulation time, was another basic measure of stroke gestures we used in this study.

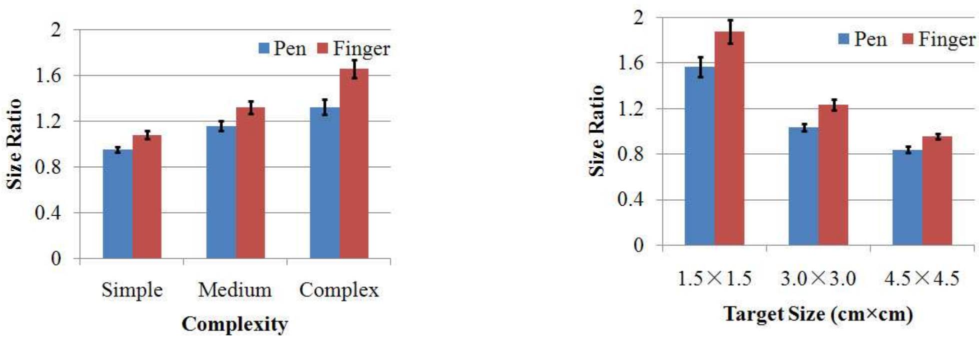
Implement had a significant main effect on average speed ($F_{1, 14} = 5.85$, $p < 0.05$). The mean speed was 6.04 cm/s for pen gestures, 6.43 cm/s for finger gestures. Complexity and target gesture size had a significant main effect on average speed ($F_{2, 28} = 59.72$, $p < 0.001$ for complexity; $F_{2, 28} = 144.78$, $p < 0.001$ for size).

Implement significantly interacted with complexity ($F_{2, 28} = 32.15$, $p < 0.001$). As shown in

Figure 3. 3a, the average speed of pen drawn gestures (6.60 cm/s) was lower than the average speed of finger drawn gestures (7.46 cm/s) in simple gestures. In addition, in medium gestures, the average speed of the pen (5.84 cm/s) was lower than the average speed of the finger (6.08 cm/s), and the average speed of the pen (5.69 cm/s) was lower than the average speed of the finger (5.76 cm/s) in complex gestures. The results indicated that the pen performed slower than the finger in the simple, medium and complex gestures, but the difference decreased from simple to complex gestures.

There was a significant interaction effect between implement and target gesture size ($F_{2,28}$ = 24.85, p < 0.001) (see Figure 3. 3b). The average speed of the pen was 4.37 cm/s, 6.20 cm/s, 7.56 cm/s for small, medium, and large target size gestures respectively while the average speed of the finger was 4.43 cm/s, 6.71 cm/s and 8.16 cm/s respectively. The finger performed faster than the pen in all three sizes, and the difference increased from small to large size.

Size Ratio



(a) (b)

Figure 3. 4 Size ratio for each implement in different (a) complexities and (b) target gesture sizes.

The participants may or may not draw the gesture exactly the same size as the target gesture displayed. There is a possibility that they would tend to draw the gesture in a larger size than the target gesture size, particularly when using the finger. The size ratio between the response gesture and the target gesture can therefore be an informative measure of the user's ability to gesture at a specified scale.

The target size (TS) of a prototype gesture has been defined in the section “Target Gesture Size,” and the response size (*RS*) of a drawn gesture is defined as the area in cm^2 of the drawn gesture's bounding box. The response to target size ratio (*SR*) was measured by the ratio of the two ($SR = RS / TS$).

Implement was a significant main effect on size ratio ($F_{1, 14} = 45.26$, $p < 0.001$). On average both pen and finger drawn gestures tended to be larger, resulting in 1.15 and 1.36 size ratio values in pen and finger conditions respectively.

The complexity had a significant effect on size ratio ($F_{2, 28} = 47.88$, $p < 0.001$). Also, there was a significant interaction effect on size ratio for gesture complexity ($F_{2, 28} = 38.64$, $p < 0.001$). As shown in Figure 3. 4a, when gesture complexity was simple, the size of drawn gestures was almost the same as the size of target gestures (mean size ratio was 0.95 for the pen, and 1.08 for the finger). Corresponding to the medium complex gestures, the mean size ratio increased to 1.16 (pen) and 1.32 (finger) respectively. For the most complex gesture, the mean size ratio increased to 1.32 (pen) and 1.66 (finger) respectively. Results showed that pen gesture led to smaller *RS* than finger gesture. In both pen and finger gestures, the size ratio increased as the gesture complexity increased.

The size ratio value strongly depended on the target gesture size ($F_{2, 28} = 71.30$, $p < 0.001$). Also, there was a significant interaction effect on size ratio for target size ($F_{2, 28} = 11.67$, $p < 0.001$). As illustrated in Figure 3. 4b, when the target size was small, the response size of the drawn gestures was larger, resulting in mean size ratio values of 1.57 (pen) and 1.88 (finger) respectively.

Corresponding to the medium size target, the mean response size of the drawn gestures was only slightly larger, resulting in mean size ratio values of 1.04 (pen) and 1.23 (finger) respectively. Corresponding to the large size target, the mean response size of the drawn gestures was in fact smaller than the size of the target, resulting in mean size ratio values of 0.84 (pen) and 0.96 (finger) respectively.

Overall, the results here show that it is difficult to draw small and complex gestures with either implement. The drawn gestures tended to be larger in these cases. These effects were slightly more pronounced with the finger than with the pen.

3.6.2 Local Shape Measures

Aperture between the Start Point and the End Point of Closed Gestures

To reflect the ability to draw a closed gesture, we measured the distance (aperture) between the start point and the end point. Conceivably the finger is at a greater disadvantage than the pen since the finger may more severely obscure the start point when getting close to it.

For drawn gestures corresponding to the prototype gestures G1, G3, G5, G7, G9 and G11 which start and end in the same position (see Table 3. 1), we calculated the aperture between the start point and the end point.

As we expected, there was a significant main effect for implement on the aperture of closed gestures ($F_{1, 14} = 5.48$, $p < 0.05$). The mean aperture was 0.20 cm with the pen and 0.24 cm with the finger respectively. Although no significant main effect was found on aperture for gesture complexity, target gesture size had a significant main effect on aperture ($F_{2, 28} = 7.86$, $p < 0.01$). The mean aperture was 0.18 cm for small size targets, 0.22 cm for medium size targets, and 0.26 cm for large size targets.

Indicative Angle Difference between Drawn Gesture and Target Gesture

The indicative angle was defined as the angle rotated from the horizontal vector whose starting point is the centroid of the gesture, to the vector formed by the centroid of the gesture and the gesture's first point. We calculated the indicative angle difference between the drawn gesture and the corresponding target gesture. It was found that no significant main effect for implement on indicative angle difference. The mean indicative angle difference was 0.32 degrees for the pen, and -0.13 degrees for the finger.

Corner Shape Distance (CSD)

(a) (b)

Figure 3. 5 (a)Vertexes in a drawn gesture. The red dots denote the vertexes. (b) The point set for each vertex. The green dots denote the point sets detected by our algorithm.

The prototype gestures G1, G2, G5, G6, G9 and G10 (see Table 3. 1) have sharp corners. How these corners change their shapes in the drawn gesture is yet another way to investigate local shape difference. We defined “Corner Shape Distance” (*CSD*) as mean distance between the corresponding corners in the drawn gesture and the target gesture.

To calculate *CSD*, as a first step we need to detect the vertex for each corner. We detected the vertexes of corners in the drawn gesture *U* based on the two-thirds power law in human motor control [52], which was also used for similar purposes in [4] to segment drawn gestures. We first calculated the speed for each point in drawn gesture *U*. Secondly, the points in *U* were sorted according to speed, and *M* (depending on the size and complexity of the corresponding target gesture) points with low speed were chosen. Third, K-means clustering was applied to partition the *M* points into *K* clusters (*K+1* was the number of corners in *U*. We did not consider the corner whose vertex is the start point, because in drawn gestures, the start point and the end point may not necessarily coincide to form a vertex.). Fourth, the point with the lowest speed in each cluster was chosen as the vertex of the corner as VC_i, $1 \le i \le K$ (see Figure 3. 5a).

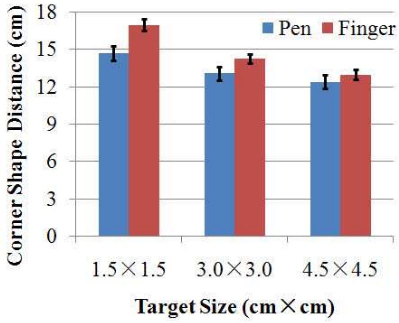


Figure 3. 6 Corner shape distance for each implement in different target sizes.

For each corner, after detecting the vertex, we need to choose a set of points in two arms to represent the corner shape. The second step is to calculate a point set for each vertex (VC_i). For each corner in the drawn gesture U, we calculated the distance between the vertex VC_i and the points in its two arms, and chose the points whose distance was less than 0.8 cm. Then each vertex (VC_i) had a point set PU_i, ($1 \le i \le K$). The points in PU_i were re-sampled into N ($N = 40$) points, which constituted a new point set $PU_{i,j}$, $(1 \le i \le K, 1 \le j \le N)$ (see Figure 3. 5b). We also calculated the point set for each vertex in the target gesture V, as PV_i, ($1 \le i \le K$), and the points in each PV_i, were also resampled into N ($N = 40$) points, which also constituted a new point set $PT_{i,j}$, $(1 \le i \le K, 1 \le j \le N)$

The third step was to calculate CSD_i ($1 \le i \le K$). CSD_i was measured by calculating the distance between the point in PD_i and the corresponding point in PT_i, ($1 \le i \le K$) [1]. To calculate the CSD_i, we translated PD_i so its centroid coincided with the centroid of the corresponding point set PT_i. The CSD was calculated by the sum of all CSD_i ($1 \le i \le K$).

$$CSD = \sum_{i=1}^{K} CSD_i = \sum_{i=1}^{K} \sum_{j=1}^{N} d(PD_{i,j}, PT_{i,j}) \tag{3.1}$$

A significant main effect for implement was found on CSD ($F_{1, 14} = 6.57$, $p < 0.05$). The mean CSD of the pen was 13.38 cm, and the mean CSD of the finger was 14.73 cm. There is also a significant main effect on CSD for target gesture size ($F_{2, 28} = 109.08$, $p < 0.01$). Implement had a significant interaction effect with target gesture size ($F_{2, 28} = 7.19$, $p < 0.01$) (see Figure 3. 6). For small target size, the mean CSD of the pen was 14.67 cm whereas the mean CSD of the finger was 16.95 cm. For medium target size, the mean CSD of the pen was 13.07 cm whereas the mean CSD of

[1] In the following sections, $d(p, q)$ was used to denote the Euclidean distance between point p and point q.

the finger was 14.25 cm. For large target size, the mean *CSD* of the pen was 12.39 cm whereas the mean *CSD* of the finger was 12.98 cm. The results showed that the pen performed better than the finger in all three target sizes.

3.6.3 Global Shape Measures

To investigate purely global shape aspects of a drawn gesture, we disregarded the drawn gesture size by normalizing (*scaling*) the drawn gesture's size to the largest target gesture size (4.5 × 4.5 cm), and also by scaling the corresponding target gesture's size to 4.5 × 4.5 cm. In other words, if the drawn gesture maintains the exact relative dimensions as the target gesture except that it is drawn in larger or smaller scale, the normalized shape measures would still give perfect scores (zero distance). We therefore report the three shape geometry measures with *scaling* (i.e. normalization). To calculate each measure, the drawn gesture was translated so its centroid coincides with the centroid of the target gesture (*translation*).

Axial Symmetry (AS)

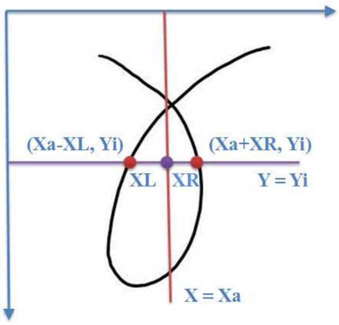


Figure 3. 7 The illustration of axial symmetry in a drawn gesture corresponding to G4.

For simplicity, we explained the algorithm of *AS* calculation by taking G4 as an example. In order to measure the drawn gesture's axial symmetry, we firstly scaled the drawn gesture to 4.5 × 4.5 cm size and then re-sampled it to *N* (*N = 500*) equidistant points. $X=X_a$ is the axis which crosses the centroid of the drawn gesture and is perpendicular to the X axis (see Figure 3. 7). For straight lines

$Y=Y_i$ ($Y_{min} \leq Y_i \leq Y_{max}$, Y_{min} and Y_{max} are the minimum y value and the maximum y value of the drawn gesture respectively, Y_i increases 1 pixel each time), there are two intersecting points between the drawn gesture and Y= Y_i: (X_a - XL, Y_i) in the left of $X= X_a$ and (X_a + XR, Y_i) in the right of $X = X_a$, in which XL is the distance between $X = X_a$ and (X_a - XL, Y_i), XR is the distance between $X = X_a$ and (X_a + XR, Y_i). The mean distance difference can be calculated as

$$AS = \frac{1}{Y_{\max} - Y_{\min}} \sum_{i=Y_{\min}}^{Y_{\max}} DA_i \tag{3.2}$$

Where DA_i is the absolute value of (XR - XL). The greater the AS is, the less symmetrical the drawn gesture is. For G5, G7, G9 and G11, the algorithm gets more complex, but the basic idea is the same.

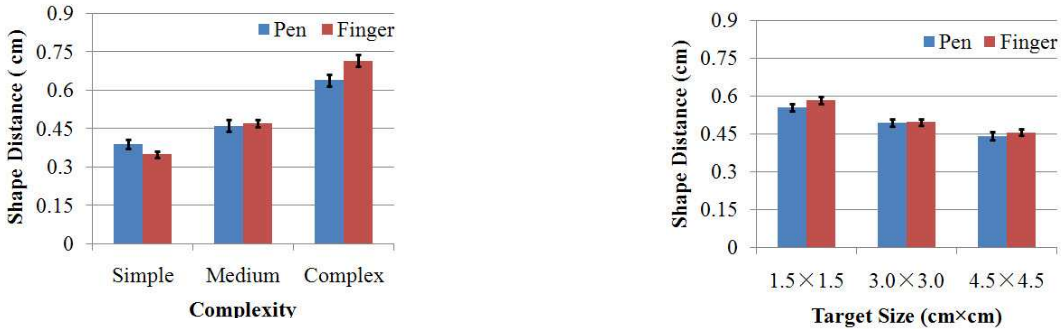
No significant main effect was found for implement on AS. The mean AS was 0.44 cm for the pen, and 0.43 cm for the finger. In other words, the finger gestures and the pen gestures did not significantly differ in symmetry. However, a significant main effect was found on AS for gesture complexity ($F_{2, 28}$ = 202.87, p < 0.001) and target gesture size ($F_{2, 28}$ = 252.89, p < 0.001). In addition, there was a significant interaction effect on AS for gesture complexity ($F_{2, 28}$ = 8.26, p < 0.01). The mean AS of pen drawn gestures (0.27 cm) was larger than that of finger drawn gestures (0.21 cm) for simple gestures, and for medium gestures, the mean AS of the pen (0.42 cm) was larger than that of the finger (0.39 cm), but the mean AS of the pen (0.62 cm) was smaller than that of the finger (0.68 cm) in complex gestures. Results showed that the finger resulted in smaller AS than the pen for simple and medium gestures, indicating that the finger performed better than the pen for these gestures.

Proportional Shape Distance (PSD)

As reviewed earlier, the *PSD* measure has gained popularity in recent years in both research and practice. Whether finger drawn and pen drawn gestures make a difference to this measure is therefore very important. After *scaling* and *translation*, the drawn gesture U and the target gesture V were re-sampled into N ($N = 100$) evenly spaced points. We denote these transformed points by $U(i)$ and $V(i)$ ($1 \leq i \leq N$) for U and V respectively.

The proportional shape distance (*PSD*) is defined as

$$PSD = \frac{1}{N}\sum_{i=1}^{N} d(U(i), V(i)) \tag{3.3}$$

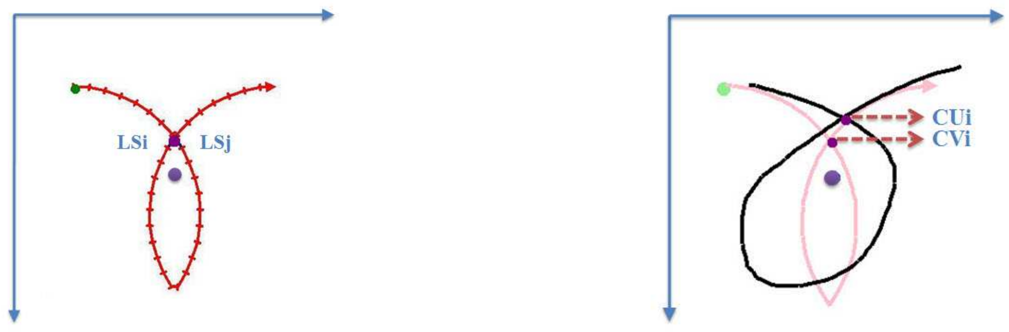


(a) (b)

Figure 3. 8 Proportional shape distance in normalized scale for each implement in different (a) complexities and (b) target gesture sizes.

Interestingly, there was no significant main effect for implement on *PSD*. The mean *PSD* was 0.50 cm for the pen, and 0.51 cm for the finger. The *PSD* measure was sensitive to both gesture complexity ($F_{2, 28}$ = 128.72, p < 0.001) and target gesture size ($F_{2, 28}$ = 150.79, p < 0.001). As one would expect, the *PSD* measure increased with gesture complexity (see Figure 3. 8a) since the accuracy to replicate more complex gestures should decrease. Furthermore, *PSD* decreased as target gesture size increased (see Figure 3. 8b). Although target gesture size had no significant interaction with implement, there was a significant interaction effect on *PSD* for gesture complexity ($F_{2, 28}$ = 7.91, p < 0.01). For simple gestures, the mean *PSD* produced by the pen (0.39 cm) was larger than that for the finger (0.35 cm). Nevertheless, for more complex gestures, the mean *PSD* achieved with the pen was smaller (0.46 cm and 0.64 cm for medium and complex gestures respectively) than that for the finger (0.47 cm and 0.72 cm for medium and complex gestures respectively). The results showed that the pen resulted in more accurate performance than the finger for complex gestures, but the finger achieved more accurate performance than the pen for simple gestures.

Intersecting Points Deviation (IPD)



(a) (b)

Figure 3. 9 (a) Detecting intersecting points in the target gesture G4. *LSi* and *LSj* are two line segments. (b) The intersecting points in G4 (*CVi*) and in the drawn gesture corresponding to G4 (*CUi*).

Gestures G2, G4, G6, G8, G10 and G12 (see Table 3. 1) had one or more self crossing intersecting points. How much these intersecting points change in the drawn gesture from the corresponding intersecting points in the target gesture is another indication of the shape difference between the two. We define the "Intersecting Points Deviation" (*IPD*) as the mean distance between the intersecting points in the drawn gesture *U* and the target gesture *V* (see Figure Figure 3. 9b).

In order to detect the intersecting points in the drawn gesture, the first step for *U* was *scaling* and *translation*, which was introduced at the start of this subsection. *U* was divided into $N-1$ ($N = 40$) line segments (LS_i, $1 \le i \le N-1$) by re-sampling into *N* equidistant points (see Figure 3. 9a). Then, the LS_i ($1 \le i \le N-1$) was compared with other line segments LS_j ($1 \le i \le N-1$, $j \neq i$) to check whether or not there were any intersecting points. If an intersecting point was detected, it would be recorded in a point set *CU*. We can also detect the intersecting points in the corresponding target gesture *V* as a point set *CV*.

If the count of intersecting points in *CU* (N_{CU}) was equal to the count in *CV* (N_{CV}), the *IPD* between *U* and *V* was calculated as

$$IPD = \frac{1}{N_{CU}} \sum_{i=1}^{N_{CU}} d(CU(i), CV(i)) \tag{3.4}$$

Else, *IPD* was calculated as 0 (*Intersection Miss*).

Intersection Miss rate, defined as the ratio of *Intersection Miss* count and total trial count for *IPD* analysis, was firstly calculated. We found that the *Intersection Miss* rate for pen input and finger input was low (3.15% and 2.09% respectively), so we continued the analysis of *IPD* using repeated measures ANOVA.

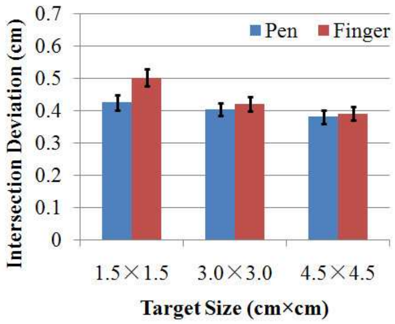


Figure 3. 10 Intersecting points deviation in normalized scale for each implement in different target gesture sizes.

A significant main effect was found for implement on *IPD* ($F_{1, 14}$ = 11.74, $p < 0.01$). Pen input resulted in *IPD* with 0.40 cm and finger input produced *IPD* with 0.44 cm. Target gesture size had significant main effects on *IPD* ($F_{2, 28}$ = 20.86, $p < 0.001$). There was a significant interaction between implement and target gesture size ($F_{2, 28}$ = 11.34, $p < 0.01$). As illustrated in Figure 3. 10, the mean *IPD* was 0.42 cm for the pen and 0.50 cm for the finger in small target size. The mean *IPD* was 0.40 cm for the pen and 0.42 cm for the finger in medium target size, and the mean *IPD* was 0.38 cm for the pen and 0.39 cm for the finger in large target size. Therefore, the pen performed better than the finger in all three target sizes.

3.6.4 Subjective Evaluation

A significant main effect was found on *speed* ($F_{1, 14}$ = 7.15, $p < 0.05$). The mean preferences of the pen and the finger were 5.53 and 4.13 respectively. A significant main effect was also found for *accuracy* ($F_{1, 14}$ = 5.59, $p < 0.05$). The mean preferences for the pen and the finger were 5.27 and 3.87 respectively. However, there was no significant main effect on *hand fatigue*, suggesting that pen gestures and finger gestures are similar in difficulty for users. Overall, users generally felt that the pen can achieve greater accuracy and faster speed than the finger for gesture input.

3.7 Discussion

3.7.1 Implications for Finger Gesture Design

Past stroke gesture research has been primarily based on the digital pen as a drawing implement. However, recent commercial product design has tended toward finger input and tends to avoid the use of the pen. Such shifts raise the question of how quantitatively different or similar finger stroke gestures are from pen stroke gestures. Therefore, we conducted a first study to quantify the differences and similarities between finger and pen gestures. Our work has provided a methodology to investigate and quantify the performance of finger and pen gestures, in which finger and pen gestures were analyzed according to multiple features that characterize stroke gestures. Some features revealed similarities between finger and pen drawn gestures, but some features were less accurate with the finger. Based on the evaluation in terms of these features, the implications for finger gesture design were presented as follows.

First, four of the nine features studied revealed similarities between finger and pen drawn gestures, including stroke articulation time (F1), indicative angle difference (F5), axial symmetry (*AS*) (F7) and proportional shape distance (*PSD*) (F8). This means that if the gesture recognition algorithm relies on features based on these measures, we should not expect finger gestures to be less effective than pen gestures. Given that proportional shape distance (*PSD*) based recognition is already used in both research and practical large-scale gesture systems (specifically the ShapeWriter gesture keyboard, although in more complex ways than in this study), it is reasonable to expect that "finger friendly" recognition algorithm can be designed within the feature space outlined by findings reported above.

Second, five of the nine features studied revealed significant differences between finger and pen drawn gestures, including the average speed (F2), size ratio (F3), aperture between start point and end point (F4), corner shape distance (*CSD*) (F6) and intersecting points deviation (*IPD*) (F9). Finger drawn gestures tended to be larger than pen drawn gestures, indicating a somewhat obvious drawback of finger operation - which requires a larger touch screen surface than pen operation. Average speed analysis revealed that the finger performed faster than the pen for gesture input, particularly for simple gestures. While the overall proportional shape distance (*PSD*) of finger gestures is no worse than pen gestures, some aspects of shape, such as intersecting points deviation (*IPD*) and corner shape distance (*CSD*) tend to be larger in finger gestures than in pen gestures. These features tend to be less accurate with the finger and thus should be avoided in “finger friendly” recognition algorithm design.

Finally, there were also a few interaction effects that may have design implications. According to time performance (F1) analysis, pen gestures led to shorter time in drawing more complex gestures. This was also reflected in movement speed (F2). The finger tended to be much faster than the pen in drawing simple gestures, but achieved similar speed in drawing complex gestures. For shape features, pen input led to more accurate axial symmetry (*AS*) (F7) than finger input for complex gestures. Furthermore, pen input is more exact than finger input for drawing complex gestures according to proportional shape distance (*PSD*) (F8), but for simple gestures, finger gestures are more accurate than pen gestures. Overall these interaction effects suggest that finger friendly gesture set design should not contain gestures which are overly complex.

All of the foregoing analysis could also be interpreted as in favor of the pen since at least in some measures it is more accurate than the finger. From daily experience in, for example hand writing and signing signatures, we can all appreciate that the dexterity of the pen is unmatchable by the finger. Note that these examples differ from the gestures tested in this experiment in at least two aspects: they tend to be more complex and they are well-learned and memorized patterns. In light of the centralist vs. peripheralist views discussed earlier in this study, one could argue that these well learned gesture patterns may include pen operation as part of one's “motor programs”.

Some interesting results were found in the subjective evaluation. In respect to speed evaluation, over half of the participants felt that pen input was faster than finger input. They stated: “The pen tip is smoother than the finger pad”. However, from average speed analysis, the finger led to higher speed than the pen for drawing gestures. Some participants reported that the finger was easier to control than the pen when drawing gestures, so they thought the finger produced higher speed than the pen. From this, we suspect that the greater degrees of freedom afforded by pen input may lead to lower drawing speed. With regard to accuracy evaluation, participants believed pen gesture input was more accurate than finger gesture input. This is consistent with the results in the analysis of corner shape distance (*CSD*) and intersecting points deviation (*IPD*), but contrary to proportional shape distance (*PSD*) analysis. Some participants reported: “It is difficult to draw intersections or sharp corners with the finger”. When drawing gestures, participants may have felt more control of some features such as corner shape distance (*CSD*) although no difference was made to other features such as proportional shape distance (*PSD*).

3.7.2 Prototype Gestures and Feature Selection

Admittedly gesture selection is a tricky balance of many considerations. We needed to cover common gestures in current usage, but we also needed to see how different types of gestures interact with various levels of complexity (simple, medium and complex) so the choices were not so many in each combination. Thus, we conducted the pilot study with a set of commonly used gestures (Graffiti gestures). Results showed pen and finger gestures differed in some features, which helped us to select features for the formal experiment. However, we did not use these Graffiti gestures in the formal experiment for two reasons. First, it is difficult to classify these gestures into simple, medium and complex levels because they are overall quite simple, i.e., these gestures can not meet the requirement of our study. Second, some features in which pen and finger gestures may differ, such as the aperture, can not be tested using these gestures because they are not closed gestures. Instead, we selected and designed twelve gestures which are suited to the purpose of this study.

Results showed that these gestures in each of simple, medium and complex gestures levels differed in terms of time performance, average speed and proportional shape distance (*PSD*), suggesting these gestures were selected properly. Furthermore, the gestures chosen for this study proved effective for our examination of the differences between finger and pen input gestures; they enabled us to reveal many plausible findings. This means that these prototype gestures may be useful also for future research when designing pen and finger gestures.

Regarding the selection of features for gesture performance measurement, although a large number of features have been proposed and used in previous studies [4], [8], [60], [80], study on all stroke gesture features is beyond the scope of our study. We only focused on some features which may reveal differences between pen and finger stroke gestures with reference to the structures and characteristics of strokes as well as the gesture input performance due to the different characteristics of pen and finger input. By means of the nine features selected or designed by us, a number of differences and similarities were found between finger and pen gestures, for example speed, size and accuracy. Furthermore, using the methodology of our study, other features can be examined for gesture design.

3.7.3 Sensing Mechanisms of the Experimental Device

The study presented here revealed that pen gesture and finger gesture differ in several features. Though we believe that the differences are caused by the intrinsic properties of the pen and the finger respectively, one may well ask whether or not the sensing qualities of different sensing mechanisms used in this study had an effect on the experimental results.

The experimental device, HP Touchsmart tx2, has two different touch sensing mechanisms, i.e., capacitive for finger input and inductive for pen input. The position accuracy and sampling rate may differ between the two sensors [22]. Regarding position accuracy, the pen tip is sharper than the finger tip, which is an inherent difference between the pen and the finger. The sampling rates in these sensors may vary depending on the number of fingers used [22]. We therefore conducted a test to measure the sampling rates of these sensors in a condition similar to our experiments. We asked all the participants to draw freely on the screen with the finger or pen. The program recorded the number of sampling points within one second. The sampling rates were 141 Hz (SD = 1.89) for pen input and 107 Hz (SD = 0.65) for finger input; both are sufficiently high for our purpose and should not affect the gesture quality measures used.

3.8 Conclusion

The rapid ongoing development of touch screen devices requires the HCI field to understand the impact of finger vs. pen as gesture implements on these devices. We conducted a first study of the differences and similarities between finger drawn and pen drawn gestures. We selected a set of gestures of varied complexity and characteristics and presented in three target sizes to a group of participants who reproduced them with both the finger and the pen. The drawn gestures were then analyzed with a broad set of measures, five selected from the literature and another four designed specifically for this study.

Our findings have demonstrated the importance of our study: when applying principles, methods and findings from pen-based gesture systems to finger-based gesture design, it is vital to consider the differences and similarities. As finger gesture interaction is gaining popularity in application design, it is important to design stroke gestures that avoid features in which the finger does not perform as well as the pen, as shown in our study. Our work is one step in this exploration.

Chapter 4 User-defined Gestures

4.1 Introduction

Information technology is increasingly being promoted as a means of support older adults to live a better life. In both personal and workplace contexts such as online banking, shopping, healthcare management, and pursuing leisure activities, older adults are a fast growing computer and Internet user group. However, perceptual, cognitive and motor deficits result in many older adults experiencing greater difficulties performing computer-related tasks than younger adults.

In order to better support older adults to interact with computers, much research has studied the changes experienced by older adults and their implications for computer use in many interactive tasks, involving three fundamental tasks in Human-computer interaction field: pointing, steering and gesturing. For a common type of motor actions in modern human computer interfaces: pointing task, a number of studies have examined the effects of aging on performance of pointing task and proposed several novel techniques to help older people perform the task [32], [34], [63]. With respect to steering task, Zhou et al. [118] and Hourcade, J.P. [31] investigated the age related effects on the performance of this kind of task.

Here, this study highlights a noteworthy interaction style: gesturing. Gesture-based interaction offers a natural and intuitive interaction style with the computer. The advantages of gesture-based interaction are two folds: (1) the user can quickly articulate the gesture for a command or a word by recalling from memory, rather than by selecting an icon with looking much at the icon; the latter is always a time-consuming process; (2) compared to tapping on the icon, gesture input is a form of more fluid movement closer to drawing, hence introducing a natural input style for the user. These two advantages make the use of gestures being gaining popularity in touch-based devices such as smart phones and Tablet PCs. Especially, due to the distinct properties of gestures, the use of gestures can help older users better interact with computers and contribute to the growth in popularity of information technology for older people. However, because younger adults are the main consumer groups of interactive devices, current gesture design aims to meet their needs, while ignores the needs of older people. Most gesture-based interfaces provided few or no accessibility features for older people, leaving the interfaces largely unusable for that age group. Although this is a serious problem, there has been very little work on the investigation of gesture performance involving the consideration of age related factors.

It is more important to design appropriate gestures for older adults than for younger adults in two reasons. First, due to the perceptual, cognitive and motor deficits of older adults, it would take longer time for them to learn how to use gestures. Hence, appropriate gestures can allow older people easy to learn and use. Second, compared to younger adults, older adults is likely to feel deeper frustration when meeting the failure to perform gestures, which hinders older adults from continuing use gestures. To better understand how older people might prefer to use gestures, we conducted a user-defined study approach that compared how older people and younger people use finger and pen gestures to perform common computing tasks on a tablet PC, so as to better design gesture-based interactions for older people.

4.2 Related Work

4.2.1 Gesture Design Based on Guessability Methodology

There are a number of gesture design studies based on guessability methodology. Earlier work by Wobbrock et al. [105] is a noteworthy analytical study in which the effects of gestures were first presented to participants and they were asked to perform actions to produce the corresponding effect. Inspired by Wobbrock's study, several works are presented. To better support the interaction between two devices including phone-to-phone, phone-to-tabletop, and phone to public display, Kray et al. [43] conducted a study in which participants were asked to spontaneously produce gestures with the mobile phone to perform a set of different activities. The study in [81] elicited motion gestures defined by participants to invoke commands on a smartphone device, hence expediting motion gesture design for the mobile computing paradigm. The study by Lee et al. [55] focused on understanding deformation-based user gestures to help in the design and implementation of deformation-based interface. Lahey et al. [53] conducted a study to understand the user of bend gestures in mobile devices with flexible electronic paper displays. Inspired by the above studies, we used a user-defined approach to examine the differences and similarities of gestures preferred by older people and younger people.

4.2.2 Pen Gesture and Finger Gesture

Pen gesture and finger gesture have been widely employed in a variety of interaction activities. We highlight some related work to show how to use pen and finger gestures in real practice.

Pen gesture have attracted a wide research interest in HCI field. For the aim of designing gesture sets that are easier to learn and more memorable, Long et al. [60] analyzed the visual similarity of pen gestures and derived a model for predicting perceived gesture similarity, based on which a gesture design tool was developed to improve the gesture design. Zhai and Kristensson [115] proposed a speed writing method, SHARK, which improved the stylus input performance with shorthand gesturing. A user study demonstrated the feasibility of the SHARK. Hinckley et al. [28] designed a ”scrapbook” application for the Microsoft Surface, in which pen and finger gesture were used to perform a number of operation, such as creating a copy of a photo by holding it and dragging off with the pen. Liao et al. [58] designed a gesture-based command system which allowed users to manipulate digital documents with paper printouts as proxies. The common operations in the daily use of computer, such as tag a paragraph, copy selections and freeform annotation can be replaced by using gestures. Appert and Zhai [5] experimentally investigated the performance and ease of learning between pen stroke shortcuts and keyboard shortcuts. They found that after adequate practice, stroke shortcuts had substantial cognitive advantages in learning and recall. Then, four guidelines about how to make stroke shortcuts easy to implement were proposed and a software prototype based on these guidelines was designed.

On the other hand, regarding to finger gestures, Wu and Balakrishnan [112] proposed a variety of multi-finger and whole hand gestural interaction techniques for multi-touch interfaces, according to which a prototype room furniture layout application, called RoomPlanner was designed. Wu et al. [113] developed and articulated a set of design principles for constructing multi-finger gestures, including gesture registration, gesture relaxation and gesture and tool reuse. Besides, a set of bimanual continuous gestures that embody these principles are developed. Morris et al. [64] described the significance of cooperative gesturing in a multi-point direct-touch surface and proposed a set of design principles with a system for collaborative art and photo manipulation, CollabDraw.

The review shows that little study paid attention to how to design gestures for older users. This study aims at shedding some light in this research field.

4.3 Experiment

Based on the methods introduced by Wobbrock et al. [105], we conducted a gesture elicitation experiment in this study. However, the study conducted by Wobbrock et al. [105] focused on touch screen interactions for younger users, the current study examines gesture preferences among both older and younger participants by means of pen input and finger input respectively.

4.3.1 Referents and Signs

A set of commands, which are application-agnostic and obtained from previous works, were used in this study. These commands are classified into two categories: (1) analogue commands; (2) abstract commands. The analogue gesture based commands can represent physical effect of real world, including move, select single, rotate, shrink, enlarge, pan, zoom in, zoom out, select group, previous, next, insert, maximize, and minimize. The abstract gesture based commands can not represent physical effect of real world, including delete, close single, duplicate, paste, undo, delete group, duplicate group, cut, menu access, and open.

4.3.2 Participants

Twenty younger participants, fifteen males and five female, aged from twenty to thirty, took part in the experiment. Ten of them were asked to design corresponding gestures with the pen according to the commands; six of the ten participants had prior experience operating digital screens with digital styli. The others were instructed to perform the same task with the finger. Nine of them had prior experience with bare finger operation on touch screen devices such as iPhone.

Twenty older participants, from sixty-five to seventy-seven of age, also took part in the experiment. None of them had the experience using the pen or the bare finger to operation on touch screen devices. Ten of them were instructed to perform the same tasks as the younger participants with the pen, while the other ten used the finger.

4.3.3 Apparatus

The study was conducted on a HP touchsmart tx2 tablet computer. The screen size of computer is 12.1 inch and its resolution was 1280 × 800 pixels. A video camera was used to record the whole experimental procedure. The experimental program was designed in the C# environment.

4.3.4 Procedure

The experimental task was similar to [105], which employed a guessability study methodology. 24 referents were randomly displayed in the screen and participants were asked to perform a gesture by means of pen input or finger input. After performing each gesture, participants were instructed to rate it on Likert Scales in terms of good match and ease to perform (7 for highest preference, and 1 for lowest preference). “Good match” refers that the gesture the participant can serve its purpose, and “ease to perform” refers that the participant can use the gesture easily.

4.4 Results

4.4.1 User-defined Pen Gesture

In this section, we analyzed pen gestures defined by older people and younger people.

Agreement Score

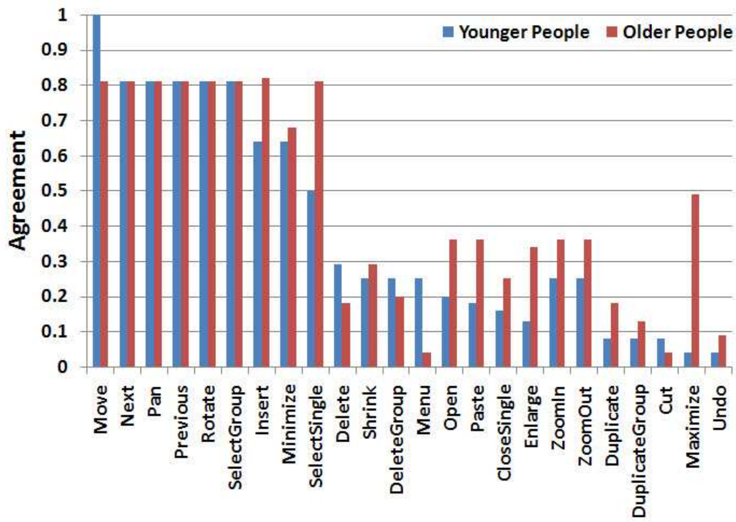


Figure 4. 1 The agreement score for all commands with pen input.

We adopted Wobbrock's method [105] to investigate the extent of agreement for each command. As an example, for the agreement of the command "Select Single", the corresponding gestures defined by 10 participants were considered. These gestures were divided into 2 groups of identical gestures. The size of each group is 5 and 5. Therefore, the agreement score for Select Single is 0.5 for younger people (Equation 1). Figure 4. 1 shows the agreement score for all commands with pen input. The overall agreement for younger people and older people was 0.39 and 0.45.

$$A=\left(\frac{5}{10}\right)^2+\left(\frac{5}{10}\right)^2 \tag{4.1}$$

As illustrated in Figure 1, some commands resulted in similar high agreement for younger people and older people, for items including *move, next, pan, previous, rotate, select group, insert* and *minimize*. For five commands: *select single, open, paste, enlarge* and *maximize,* younger people and older people had different agreement scores.

Good Match and Easy to Perform

For all commands which had similar user-defined gestures with pen input and finger input, T statistic showed no significant difference on good match and on easy to perform between younger people and older people ($p > 0.05$). While for some commands which led to different user-defined gestures: *open* and *paste*, a significant different was found on good match and on easy to perform between younger people and older people ($p < 0.05$); younger people rated higher than older people. For other commands: *select single, enlarge* and *maximize,* no significant different was found on good match and on easy to perform between older people and younger people regarding pen input and finger input ($p > 0.05$). Overall, older and younger people rated “analogue gesture based command” higher than “abstract gesture based command”, indicating that analogue gestures were easier than abstract gestures for users to remember and perform.

4.4.2 User-defined Finger Gesture

In this section, we analyzed pen gestures defined by older people and younger people.

Agreement Score

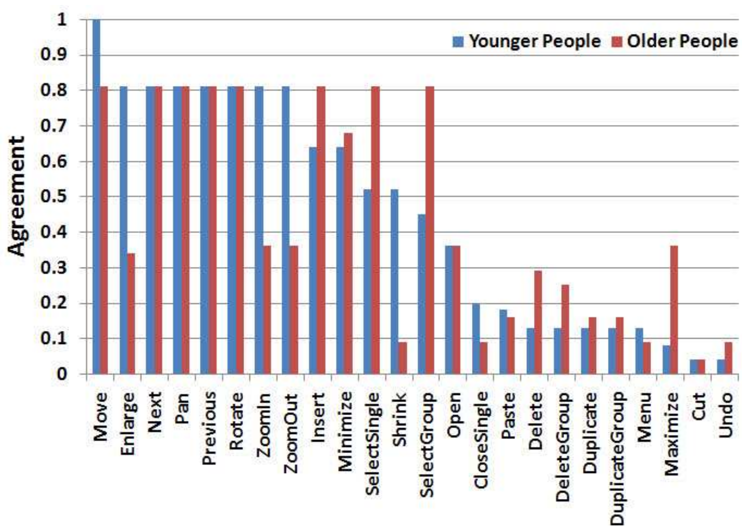


Figure 4. 2 The agreement score for all commands with finger input.

How to calculate agreement score is described in previous section. Figure 4. 2 shows the agreement score for all commands with finger input. The overall agreement for younger people and older people was 0.46 and 0.43. *Move, next, pan, previous, rotate, insert* and *minimize* produced similar high agreement score, while *enlarge, zoom in, zoom out, select single, select group,* and *maximize* led to different agreement score.

Good Match and Easy to Perform

For all commands which had similar user-defined gestures with pen input and finger input, T statistic showed no significant difference on good match and on ease to perform between younger people and older people ($p > 0.05$). While for some commands which led to different user-defined gestures: *enlarge, zoom in, zoom out* and *maximize*, a significant difference was found on good match and on ease to perform between younger people and older people ($p < 0.05$); younger people rated higher than older people. For other gestures: *select single and select group*, no significant different was found on good match and on easy to perform between younger people and older people ($p > 0.05$). Overall, users rated "analogue gesture based command" higher than "abstract gesture based command", indicating that analogue gestures were easier than abstract gestures for users to remember and perform.

Preference for Number of Hands

In the experiment, the participants were allowed to perform gestures with multi-fingers. However, when performing gestures to invoke commands, younger people preferred 1-hand gestures for 25 of 27 referents, while older people commonly used only one finger. This may be due to the fact that younger people were familiar with multi-touch input but older people were not. Using one finger is a preferred approach to perform gestures for older people.

4.5 Discussion

4.5.1 Analogue and Abstract commands

According to the result analysis, for both pen input and finger input, younger people and older people all preferred analogue gestures. This may be due to the fact that analogue gesture can represent physical effect of real world. Hence, we suggest that when designing gestures for older people, analogue gestures should be deeply explored and widely used.

4.5.2 Desktop Paradigm for Gesture Design

Although we took care not to show elements from Windows or the Macintosh, younger participants still often thought of the desktop paradigm. For example, they used two kinds of gesture to select single target: tap on an object or draw a circle around an object. However, because older participants seldom used computers, they are not familiar with the desktop paradigm. Therefore, they commonly drew a circle around the target to perform selection. The results indicated that gesture design for younger and older people should consider the effect of the desktop paradigm.

4.5.3 Multi-touch Gesture for Older People

Because older people seldom use multi-touch devices, they did not know how to leverage gesture input with multi-touch property. In addition, due to the motor deficit, the older people may feel difficult to perform multi-touch gesture. This indicates that in order to employ multi-touch gesture for older people, it is important to educate the older people how to use multi-touch input.

4.5.4 Finger Gesture and Pen Gesture

For younger people and older people, finger gesture and pen gesture differ in some commands which need high DOF to perform: shrink, enlarge, zoom in, and zoom out. Pen gesture design should avoid using high DOF degree gestures for older and younger users.

4.6 Conclusion

We examined and compared user-defined gestures between younger people and older people. To design better gesture for older people, it was found that 1) gesture design should avoid using gestures with high Degree of Freedom; 2) desktop paradigm has less effect on gesture performance for old people than for younger people; 3) analogue gestures should be deeply explored and widely used for older people. The results can be beneficial to gesture interface design for elder people and for younger people.

Chapter 5 Optimal Entry Size of Handwritten Chinese Characters in Touch-based Mobile Phones

5.1 Introduction

Touch-based mobile phones have received much attention in recent years, for they allow users to directly manipulate digital information using fingers instead of a pen or keyboard. For mobile phones such as iPhone, which does not have a physical keyboard, one of the commonly used character entry styles is handwritten character entry. This entry style is important in script languages such as Chinese, Japanese and other non-alphabetic languages [83]. However, the small screen area of mobile phones restricts the handwriting entry area size. In order to design a rational screen layout which can display more information and also allow users to write with ease and high efficiency, it is important to determine the optimal entry area for the handwriting of characters. For example, when a user edits a text file in a mobile phone, it is desirable that the screen display as many characters as possible and that the probability of the user having to drag the scrollbar to view text information be as low as possible. Although some devices provide unframed handwriting interfaces, the optimal entry area parameter is still important for determining the display area of the possible recognized characters corresponding to an inputted character; a larger area can support the designer to design a more suitable display style for the recognized characters [78] and can allow the user to select a recognized character more easily with a finger [72].

The optimal size of handwriting character input boxes for a stylus on PDAs has already been investigated by [79]. The optimal entry size in that study was defined as the smallest input area in which the user can input characters with short writing time, small number of error corrections, small number of stroke protrusions outside the area and high subjective assessment (for example, ease of writing and degree of fatigue). Ren and Zhou [79] considered entry boxes for different kinds of characters, different box sizes and shapes, different user postures and different user age groups. Four dependent variables, including number of protruding strokes, number of error corrections, writing time and subjective preference of participants, were used to assess handwriting performance for each entry box size. The optimal size of an entry box for the input of alphanumeric characters was found to be 1.2 × 1.4 cm (rectangular), whereas for kanji (Chinese characters) mixed with kana characters and for hiragana and katakana characters, the optimal size of an entry box was found to be 1.4 × 1.4 cm (square). However, finger input is less precise than stylus input, and writing characters using a pen is more familiar for users than writing with a finger. Therefore, the conclusions drawn by Ren and Zhou may not apply to touch-based mobile phones. There remains a need for investigation of optimal finger-based character entry size in touch-based mobile phones.

In this study, we define the optimal entry size for character handwriting as the smallest input area in which the user can input characters with high entry area utilization rate, great writing speed, high character recognition rates, small number and short length of stroke protrusions outside the area and high subjective assessment (for example, ease of writing and degree of fatigue). For touch-based mobile phones, two commonly used handwriting styles with fingers are 1) two-handed entry with the non-dominant hand holding the device and the index finger of the dominant hand entering characters, usually with the user sitting; 2) one-handed entry with the dominant hand holding the device and the thumb of the dominant hand entering characters, usually with the user walking. This study focuses on determining the optimal entry size of handwritten characters through two experiments; one to investigate the optimal entry size for two-handed entry, and the other to examine the optimal entry size for one-handed entry. To determine the optimal entry box size, we defined a set of dependent variables for handwriting performance measures and proposed a variation of an existing experimental paradigm.

In the following sections, related work on handwriting input is firstly described. This is followed by a description of experiment design, including the selection of entry box sizes, determination of entry box position, definition of dependent variables for performance measures and a report on experimental devices. Two experiments are then reported, and after each experiment, experimental results are discussed. Then we present a general discussion of the results and discuss future work, after which a conclusion is finally drawn.

5.2 Related Work

This work builds upon three distinct areas of previous research. The first refers to the design of handwriting entry boxes. The second is a body of work on the improvement of handwriting performance. The last is two-handed and one-handed use of touch-based mobile phones. We review each in turn.

5.2.1 Handwriting Entry Box Design

This section reviews handwriting entry box design from aspects of fundamental research and commercial product design. In a noteworthy fundamental study, Ren and Zhou [79] compared different entry box sizes and shapes for pen-based handwriting on PDAs and found that the optimal size for an entry box for the input of alphanumeric characters is 1.2 × 1.4 cm (rectangular), whereas for kanji (Chinese characters) mixed with kana characters and for hiragana & katakana characters the optimal size is 1.4 × 1.4 cm (square). On the basis of that work, we conducted a series of experiments to determine the optimal handwriting entry box size in touch-based mobile phones. On the other hand, for the common mobile operating systems (OS) used by modern touch-based mobile phones, such as Microsoft's Windows Mobile and Windows Phone (Microsoft Corp.), Apple's iOS (Apple Inc.) and Google's Android (Google Inc.), the position and size of the handwriting entry box may differ according to software applications supported by these OSs. Taking Apple iPhone 4 (Apple Inc.) as an example, the handwriting entry box is set at the bottom of the screen with a size of 3.3cm × 4cm. However, it remains unclear whether this entry area can support fast, accurate and ease of handwriting with high utilization rate of entry area.

5.2.2 Improvement of Handwriting Performance

The aim of this study is to determine the optimal entry box area within which users can handwrite characters quickly and accurately. We review some studies that have paid attention to the analysis of handwriting speed and accuracy, and also to the improvement of handwriting speed and the reduction of handwriting errors.

In a noteworthy analytical study, MacKenzie et al. [62] experimentally analyzed three character entry methods for pen-based computers, with evaluation in terms of entry speed and accuracy, for aspects including handwriting input, tapping on a soft keyboard with a QWERTY layout and tapping on a soft keyboard with an ABC layout. Handwriting produced a writing speed of 16 wpm, slower than the writing speed for tapping on a QWERTY soft keyboard but quicker than the writing speed produced by tapping on the ABC soft keyboard; handwriting led to 8.1% entry errors, which is greater than the error rate of the other two techniques. Commarford [20] compared the usability of Graffiti and a virtual keyboard on a PDA running Microsoft Windows CE and found that participants performed better with the virtual keyboard but showed no preference for the program. On the other hand, to investigate unconstrained handwriting performance, Kristensson and Denby [44] conducted an experimental study based on unconstrained handwriting recognition. In the recognition method, the recognizer simultaneously accepted hand-printed characters and cursive script. The mean difference in entry rate and error rate between software keyboard and unconstrained handwriting recognition was not significant, which indicates that performance in the two entry techniques is similar.

There have also been some studies of speed enhancement and error reduction. Wobbrock et al. [106] proposed a word-level stroking system, which aims to improve the speed of character-level unistrokes. Kurihara et al. [48] proposed a multimodal input system that can provide multiple prediction characters, enabling greater handwriting speed and fewer handwriting errors. That system provides multiple predictions based on speech recognition and handwriting recognition, and the user selects one item and pastes it in the edit board avoiding tedious manual writing. For correcting handwriting errors, Shilman et al. [92] proposed a mixed-initiative approach, which can continually evolve the recognizer's results using the additional information from user corrections. A user study demonstrated the effectiveness of this error correction approach. Various handwriting methods have also been proposed for computer-based speed writing. As a well known single stroke shorthand handwriting recognition, Graffiti (Palm Computing, Inc) has been widely used in PDAs based on the

Palm OS. Unistrokes alphabet is a gesture alphabet for stylus-based text entry [25], in which every letter is written with a single stroke but the more frequent ones are assigned to simpler strokes. Shapewriter [115], a novel form of writing that uses pen strokes on graphical keyboards to write text, can enable users to enter text efficiently at a faster rate than previously possible on mobile phones, handheld computers and other mobile devices. Wobbrock et al. [107] proposed EdgeWrite, a new unistroke text entry method for handheld devices designed to provide high accuracy and stability of motion for people with motor impairments.

5.2.3 Two-handed and One-handed Use of Touch-based Mobile Phones

Two-handed and one-handed use of touch-based mobile phones has generated considerable recent research interest. We focused on the studies which investigated appropriate target size for two-handed and one-handed use, because these studies inspired our approach of examining optimal entry box size.

For a data entry task on a PDA with a stylus, Sears and Zha [89] investigated the effect of soft keyboard size (small, medium and large) for two kinds of soft keyboard. According to the analysis on three measures, data entry rate, uncorrected error rates and subject preferences, they drew the conclusion that keyboard size does not affect data entry rates, error rates and preference ratings. However, for text entry on a desktop-sized touch screen with a finger, Sears et al. [88] found larger key size text resulted in higher entry rates for both novice and experienced users, and that novices committed significantly fewer errors on the largest keyboard than on the smallest one.

Regarding one-handed use, Karlson et al. [38] conducted a systematical study to understand single-handed mobile device interaction. The study revealed that users can tap faster in the center area of mobile phone screen than other screen regions for small candy bar phones, flip phones, large candy bar phones and PDAs. According to this finding, we set the entry box position in the center area of the mobile phone screen in our experiment. Parhi, et al. [72] investigated target size for one-handed thumb use on small touch screen devices. With analysis in terms of task time, error rate, hit distribution, and user preference, they found that a target size of 9.2 mm for single-target pointing and a target size of 9.6 mm for a sequence of taps should be sufficiently large without degrading performance and preference. They also found that users can tap faster in the center area of mobile phone screen than in other screen regions, which is consistent with Karlson et al.'s finding.

Our review indicates that little study has been undertaken to determine optimal handwriting entry box size for two-handed use as well as one-handed use for touch-based mobile phones. Therefore, we set out to perform a systematic investigation of optimal finger-based handwriting entry box size, so as to provide guidelines on user interface design of handwriting.

5.3 Experiment Design

In our experiment, an entry box was shown in the screen of the experimental mobile phone, and experimental participants were asked to write the corresponding character within the entry box according to a prototype character which was displayed in the mobile phone. Selection of entry box size, determination of entry box position and shape, and selection of prototype characters will be described in the following three subsections, after which a set of dependent variables for performance measures and configurations of experimental devices are detailed.

5.3.1 Entry Box Size (EBS)

Five levels of entry box sizes, 1.5cm×1.5cm, 2.0cm×2.0cm, 2.5cm×2.5cm, 3.0cm×3.0cm and 3.5cm×3.5cm were used in our experiment. The setting of the minimum entry box size at 1.5cm×1.5cm is based on the results obtained in [79] which suggested that the optimal size of a pen-input box for a stylus-based PDA was 1.09cm×1.66cm- 1.44cm×1.44cm for alphanumeric characters and 1.44cm×1.44cm for kanji and kana. We suspected that 1.5cm×1.5cm size would be more challenging for finger input than for stylus input. In the following sections, EBS (1.5 × 1.5), EBS (2.0 × 2.0), EBS (2.5 × 2.5), EBS (3.0 × 3.0) and EBS (3.5 × 3.5) are used to denote the entry box size with 1.5cm×1.5cm, 2.0cm×2.0cm, 2.5cm×2.5cm, 3.0cm×3.0cm and 3.5cm×3.5cm respectively.

5.3.2 Entry Box Position and Shape

Because our study focuses on determining the optimal entry box size for handwriting input, the entry box should be set in a position in which users can perform handwriting input quickly and easily. Although entry box position may not influence two-handed entry performance, it does affect one-handed entry performance. The study conducted by Karlson et al. [38] showed that thumb interaction performance on the surface of mobile phones is related to the screen region; different regions result in different tap speeds. For flip phones, the center of the screen was found to be able to support high tap speed. In our experiment, therefore, the entry box was always shown in the center of the screen of the experimental device. The entry box shape was set as square due to the fact that the bounding box of almost all Chinese characters is square.

To examine one-handed handwriting performance in a center entry box with 2.5cm × 2.5cm, we conducted a preliminary experiment with four people, two males and two females. In the experiment, the participants were asked to write ten arbitrary Chinese characters within the entry box. All participants stated that the writing task was easy to perform. Therefore, we believe a square entry box placed in the center can support fast and easy handwriting input.

5.3.3 Prototype Character Categories

Character Complexity	Stroke Number	Left to Right	Top to Bottom	Mixed
Simple	2 strokes	八	丁	刀
	3 strokes	小	于	女
	4 strokes	以	午	区
Medium	7 strokes	汽	来	困
	8 strokes	和	学	国
	9strokes	指	是	逆
Complex	12strokes	湖	番	属
	13 strokes	靴	罪	遥
	14 strokes	歌	算	遮

Table 5. 1 The prototype Chinese characters categories.

In order to investigate character entry performance among the five entry box sizes, we selected 27 commonly used Chinese characters (see Table 5. 1), which were divided into three groups (simple, medium, complex) according to the number of strokes making up the character. The structures of these characters were divided into left to right, top to bottom and mixed structure respectively.

5.3.4 Performance Measures

Handwriting Speed	Utilization Rate of Entry Area	Accuracy	Ease of Writing	Subjective Evaluation
1. Writing time 2. Stroke writing speed	1. Size ratio	1. Number of writing attempts	1. Number of protruding strokes 2. Length of protruding strokes	1. Preference Ratings on writing speed, utilization rate of entry area, writing accuracy, ease of writing and finger fatigue

Table 5. 2 The classification of dependent variables.

As mentioned in the section “Introduction”, Ren and Zhou [79] used four dependent variables to determine the optimal handwriting character input box size for stylus on PDAs. Referring to these four dependent variables, we defined a set of dependent variables and sorted these variables into five aspects illustrated in Table 5. 2. The five aspects were handwriting speed, utilization rate of entry area, accuracy, ease of writing and subjective evaluation.

Writing Time

The writing time is defined as the time duration from the moment the finger touches the screen to the moment the last stroke is finished. It should be noted that writing time involves the time interval from the end of a stroke to the start of the subsequent stroke. This performance measure describes the overall time of the writing procedure.

Stroke Writing Speed

The stroke writing speed is calculated by the ratio of an inputted character's length and the corresponding stroke writing time which is defined as the length of time (duration) that the finger touches the screen during handwriting input.

Size Ratio

The participants may or may not write the character exactly the same size as the entry area. There is a possibility that they would tend to write the character within an area which is smaller or larger than the entry area. To examine the utilization rate of entry area, we measured “size ratio”, defined as the ratio of the size of the bounding box of the written character and the size of entry box. A higher size ratio indicates that the user utilizes a larger area for an entry box for handwriting input.

Length and Number of Protruding Strokes

If a part of a writing stroke is outside an entry box, a protruding stroke is detected. The number and length of protruding strokes are recorded respectively to describe the difficulty of inputting handwriting within the entry box. If the user inputs a character easily within a given area, there should be only short length and a small number of protruding strokes outside of the defined input area.

Number of writing attempts

The number of writing attempts is used to describe handwriting accuracy. Microsoft Windows XP Tablet PC Edition 2005 Recognizer software (Microsoft Corp.) which can recognize Chinese and Japanese handwriting, was applied to written character recognition. If an inputted character can be recognized as the corresponding prototype character, the current writing trial will finish and the subsequent writing trial will start; otherwise, the number of writing attempts increased by one and participants were asked to input the character again. Our approach for handwritten character recognition is totally different from Ren and Zhou’s approach (2009) in which the recognition result relied on participants’ subjective judgment.

Subjective Preference

Writing performance for each entry box was rated by participants according to five dimensions: writing speed, utilization rate of entry area, ease of writing, writing accuracy and fatigue of the finger used for writing. Participants were required to rate these entry boxes on a 5-point scale (1 for worst, and 5 for best).

5.3.5 Experimental Device

This study was conducted on a HTC Touch HD mini smartphone for handwriting input, and a 1.66 GHz Intel Core2 PC with Windows XP Professional SP2 for handwritten character recognition. The smart phone has a capacitive touch screen with HVGA resolution; the screen size is 3.2 inch and its resolution is 320 × 480 pixels. The smartphone platform is Windows Mobile 6.5 Professional with HTC Sense. The smartphone was connected to the PC via Wi-Fi networks. In both devices, experimental programs were designed in the C# Environment.

5.4 Experiment One (Two-Handed Entry)

5.4.1 Participants

Nine participants, 2 males and 7 females, from 20 to 32 years of age, participated in this experiment. Three of them were Japanese and the others were Chinese. All of them were right-handed and had prior experience with bare finger operation on touch screen devices such as an iPhone. The physical sizes of each subject’s finger-tips (end joints) were recorded. The average values of physical width (W) and physical length (L) were listed in Table 5. 3.

	Thumb		Index Finger	
Finger Tip	**L**	**W**	**L**	**W**
AVG	**31.2**	**21.2**	**25.4**	**16.3**
SD	**2.8**	**2.9**	**2.0**	**2.1**

Table 5. 3 The average values (millimeter) of physical width (W) and physical length (L) of the thumb and the index finger.

5.4.2 Task and Procedure

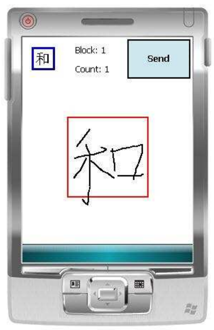


Figure 5. 2 Experimental interface.

Figure 5. 1 User in the two-handed entry environment.

When performing the experimental task, participants were asked to sit in a chair and hold the device with the non dominant hand. In each test trial, a prototype character was shown in the top left corner (see Figure 5. 2), and participants were asked to write a corresponding character using the index finger of the dominant hand within the red box as quickly and clearly as possible (see Figure 5. 1). After finishing writing the character, participants were instructed to press the send button so that information about the inputted character could be sent to a PC in which the handwriting recognition software was running. The character recognition process was detailed in the section "Number of writing attempts". Each participant completed 2 blocks of 27 prototype characters in 5 sizes. Within each block, the order of the 27 characters in 5 different sizes was randomized. In summary, experiment data collection consisted of:

9 subjects ×

2 blocks of trials ×

27 characters ×

5 target entry sizes

= 2430 drawing trials

At the end of the experiment, a questionnaire was administrated to gather subjective opinions.

5.4.3 Results and Analysis

Writing Time (WT)

Repeated measures ANOVA showed that EBS had no significant main effect on writing time ($F_{4,32}$ = 0.592, p = 0.671). Post-hoc comparisons using the Bonferroni adjustment for multiple comparisons found no significant difference between all EBSs. For EBS (1.5 × 1.5), EBS (2.0 × 2.0), EBS (2.5 × 2.5), EBS (3.0 × 3.0) and EBS (3.5 × 3.5), the mean WT was 2497ms, 2532ms, 2523ms, 2532ms and 2557ms respectively. There was no interaction effect on WT for character complexity × EBS ($F_{8,64}$ = 0.42, p = 0.908).

Stroke Writing Speed (SWS)

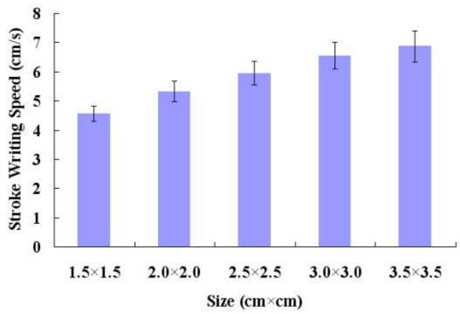


Figure 5. 3 The mean stroke writing speed for each entry box size with two-handed entry. Error bars represent 0.95 confidence interval.

Repeated measures ANOVA showed a significant main effect on SWS for EBS ($F_{4,32}$ = 46.284, p < 0.001) (see Figure 5. 3). Also, there was an interaction effect on SWS for character complexity × EBS ($F_{8,64}$ = 4.432, p < 0.001). Post-hoc comparisons were performed using the Bonferroni adjustment for multiple comparisons, and it found no significant difference between EBS (2.5 × 2.5) (Mean = 5.948 cm/s) and EBS (3.5 × 3.5) (p = 0.010) (Mean = 6.873cm/s), EBS (3.0 × 3.0) (Mean = 6.543cm/s) and EBS (3.5 × 3.5) (p = 0.129). Although these three EBSs can be grouped on similar SWSs, they were significantly faster than EBS (1.5 × 1.5) (p < 0.005) (Mean = 4.560cm/s) and EBS (2.0 × 2.0) (p < 0.005) (Mean = 5.327cm/s).

Size Ratio (SR)

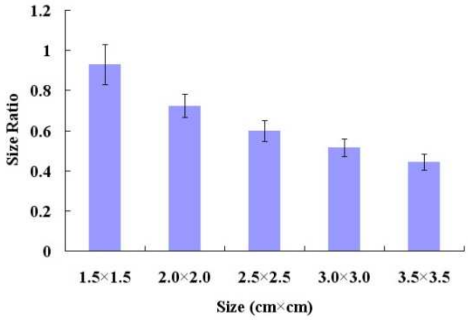


Figure 5. 4 The mean size ratio for each entry box size with two-handed entry.

Repeated measures ANOVA showed that EBS had a significant main effect on SR ($F_{4,32} = 40.912$, $p < 0.001$). As illustrated in Figure 5. 4, smaller EBS usually led to larger SR. For EBS (1.5 × 1.5), EBS (2.0 × 2.0), EBS (2.5 × 2.5), EBS (3.0 × 3.0) and EBS (3.5 × 3.5), the mean SR was 0.929, 0.723, 0.598, 0.514 and 0.443 respectively. There was also a significant interaction between character complexity and EBS ($F_{8,64} = 27.703$, $p < 0.001$). Post-hoc comparisons using the Bonferroni adjustment showed significant difference ($p < 0.005$) between all EBSs except EBS (1.5 × 1.5) and EBS (2.0 × 2.0) ($p = 0.023$). Therefore, these two EBSs can be grouped on their similar SRs.

Length of Protruding Strokes (LPS)

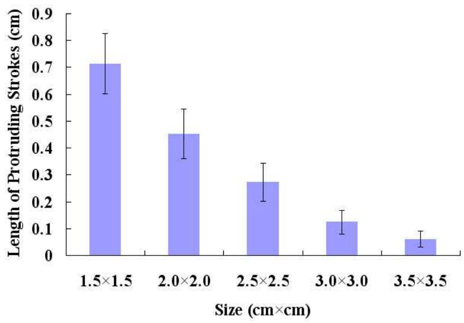


Figure 5. 5 The mean length of protruding strokes for each entry box size with two-handed entry.

An analysis of normality found that data of LPS were skewed, so a square root transformation was applied to remedy the data. Repeated measures ANOVA showed that EBS had a significant main effect on LPS ($F_{4,32}$ = 48.682, p < 0.001). Figure 5 shows that LPS tended to be reduced when EBS increased. Also, there was an interaction effect on LPS for character complexity × EBS ($F_{8,64}$ = 39.982, p < 0.001). Post-hoc comparisons were performed using the Bonferroni adjustment for multiple comparisons. The LPS of EBS (3.5 × 3.5) was the smallest with a mean of 0.061cm, which showed no significant difference (p = 0.123) from EBS (3.0 × 3.0) (Mean = 0.124cm). No significant difference was found on LPS between EBS (3.0 × 3.0) and EBS (2.5 × 2.5) (Mean = 0.273cm). Moreover, the LPSs produced by these three EBSs are significantly smaller than the LPSs produced by EBS (1.5 × 1.5) (Mean = 0.714cm) (p < 0.005) and EBS (2.0 × 2.0) (Mean = 0.453cm) (p < 0.005). Therefore, EBS (2.5 × 2.5), EBS (3.0 × 3.0) and EBS (3.5 × 3.5) can be grouped on their similar LPSs.

Number of Protruding Strokes (NPS)

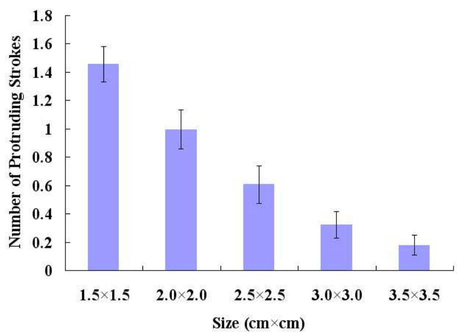


Figure 5. 6 The mean number of protruding strokes for each entry box size with two-handed entry.

Because an analysis of normality found that the data of NPS were skewed, a square root transformation was used to remedy the data. Repeated measures ANOVA showed that EBS had a significant main effect on NPS ($F_{4,32}$ = 146.438, p < 0.001). As illustrated in Figure 5. 6, smaller EBS led to larger NPS. For EBS (1.5 × 1.5), EBS (2.0 × 2.0), EBS (2.5 × 2.5), EBS (3.0 × 3.0) and EBS (3.5 × 3.5), the mean NPS was 1.457, 0.996, 0.608, 0.323 and 0.178 respectively. There was also an interaction effect on NPS for character complexity × EBS ($F_{8,64}$ = 29.761, p < 0.001). Post-hoc comparisons using the Bonferroni adjustment showed a significant difference (p < 0.005) between all EBSs except EBS (3.0 ×3.0) and EBS (3.5 × 3.5) (p = 0.030). Therefore, EBS (3.0 ×3.0) and EBS (3.5 × 3.5) can be grouped on their similar NPSs.

Number of Writing Attempt (NWA)

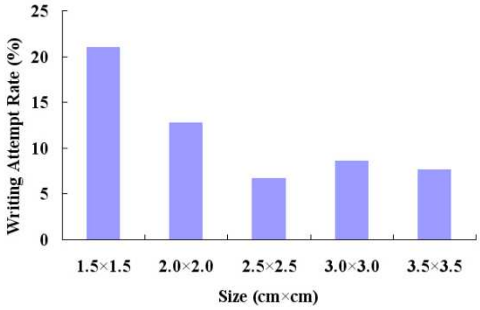


Figure 5. 7 The writing attempt ratio for each entry box size with two-handed entry.

Data of NWA was analyzed by using the chi-square test based on standardized residuals. On the frequency of NWA, we found that there is a statistically significant relationship between NWA and EBS (χ^2_4 = 11.31, p < 0.05) (see Figure 5. 7). Interestingly, EBS (2.5 × 2.5) seemed to support the most accurate handwriting input, because it resulted in the fewest NWA with a standardized residual at -4.6. EBS (3.5 × 3.5) and EBS (3.0 × 3.0) produced almost the same NWAs (z = -3.6 and z = -2.6 respectively) as EBS (2.5 × 2.5). The largest NWA was produced by EBS (1.5 × 1.5) (z = 9.4).

Subjective Preference

Dimension	EBS (1.5 × 1.5)	EBS (2.0 ×2.0)	EBS (2.5 × 2.5)	EBS (3.0 ×3.0)	EBS (3.5 × 3.5)
Writing speed	1.4	2.2	3.4	4.6	4.7
Utilization rate	4.9	4.6	4.1	3.4	2.8
Writing accuracy	1.4	2.2	3.9	4.3	4.8
Ease of writing	1.2	2.3	3.4	4.7	4.9
Fatigue of the finger	1.4	2.1	3.6	4.4	4.7

Table 5. 4 The participants' preferences for each EBS with two-handed entry.

There was a significant difference found in the effect of the five EBSs on the overall preference rating ($F_{4,32}$ = 53.059, p < 0.001) (see Table 5. 4). Post-hoc comparisons using the Bonferroni adjustment showed that EBS (3.5 × 3.5), EBS (3.0 × 3.0) and EBS (2.5 × 2.5) were rated significantly higher than the other two EBSs (p < 0.005). Moreover, no significant difference was found between EBS (3.5 × 3.5), EBS (3.0 × 3.0) and EBS (2.5 × 2.5). Therefore, these three EBSs can be grouped on the similar subjective preference.

Learning Effect

Repeated measures ANOVA showed no obvious learning effect on the WT ($F_{1,8}$ = 2.028, p = 0.192), SRS ($F_{1,8}$ = 3.305, p = 0.107), LPS ($F_{1,8}$ = 0.151, p = 0.707) and NPS ($F_{1,8}$ = 0.146, p = 0.713) between the two experimental blocks. As well, no learning effect was found on the frequency of NWA (χ_1^2 = 0.01, p = 0.948) using a chi-square analysis. An important reason that the participants could keep their handwriting performance stable is that they were familiar with the prototype Chinese characters.

5.4.4 Discussion

Handwriting Speed

Writing time (WT) and Stroke writing speed (SWS) is used to quantify handwriting speed. Interestingly, our results show that for the five EBSs, users spent similar writing time on the handwriting task. However, the analysis on SWS reveals that different EBSs can lead to different stroke writing speeds. According to the analysis results, EBS (2.5 × 2.5) is regarded as the smallest EBS in which users can perform handwriting task with high speed.

Utilization Rate of Entry Area

It is evident that users prefer a large entry area for handwriting input. However, a larger entry area may lead to a lower utilization rate of entry area; a handwriting character may occupy only a small area. In this study, size ratio (SR) is applied to quantification of entry area utilization rate; larger SR represent higher utilization rate. Our results indicate that EBS (1.5 × 1.5) and EBS (2.0 × 2.0) resulted in a similar high utilization rate of entry area. Although the SR produced by EBS (2.5 × 2.5) was smaller than that produced by EBS (1.5 × 1.5) and EBS (2.0 × 2.0), it was significantly larger than that produced by EBS (3.0 × 3.0) and EBS (3.5 × 3.5), suggesting that EBS (2.5 × 2.5) could also generate a high utilization rate of entry area.

Ease of Writing

Length and number of protruding strokes (LPS and NPS) are applied to the description of the ease of writing; shorter LPS and smaller NPS represent easier handwriting input within a given area. Although the NPS produced by EBS (2.5 × 2.5) was larger than that produced by EBS (3.0 × 3.0) and EBS (3.5 × 3.5), these three EBSs resulted in similar LPSs. Therefore, EBS (2.5 × 2.5) can be considered to be the smallest entry box size in which users can perform a handwriting task easily.

Entry Accuracy

The number of writing attempts (NWA) is used to describe entry accuracy. Although the entry accuracy mainly depends on the recognition ability of handwriting recognition software, our results indicate that entry box size had an effect on entry accuracy. For EBS (1.5 × 1.5), the reason why this size produced the lowest entry accuracy may be that users can not write each stroke clearly within the small area. EBS (2.5 × 2.5) seems to support the most accurate handwriting input, because it resulted in the fewest NWA. Moreover, EBS (3.0 × 3.0) and EBS (3.5 × 3.5) led to similar entry accuracy as that produced by EBS (2.5 × 2.5). Therefore, we believe EBS (2.5 × 2.5) is the smallest entry box size in which users can perform handwriting input with high entry accuracy.

Subjective Preference

According to the preference ratings of these five EBSs, EBS (2.5 × 2.5) had a comparatively high overall rating. The result is fairly consistent with the analysis of handwriting speed, utilization rate of entry area, ease of writing and entry accuracy.

In summary, for two-handed handwriting input, the optimal entry box size was found to be 2.5cm × 2.5cm.

5.5 Experiment Two (One-Handed Entry)

5.5.1 Participants and Equipment

The same nine subjects who participated in the Experiment one took part in Experiment two. The same equipments were used as in Experiment one.

5.5.2 Task and Procedure

Figure 5. 8 User in the one-handed entry environment.

A similar writing task and procedure used in Experiment one was carried out in this experiment. The only difference was that in this experiment participants were asked to perform the writing task with the thumb of their dominant hand while holding the device with the dominant hand (see Figure 5. 8). The mean physical width and mean physical length of the thumb used for writing were listed in Table 3.

5.5.3 Results and Analysis

Writing Time (WT)

Repeated measures ANOVA showed that EBS had no significant main effect on writing time ($F_{4,32}$ = 0.805, p = 0.531). Post-hoc comparisons using the Bonferroni adjustment for multiple comparisons revealed no significant difference between all EBSs. For EBS (1.5 × 1.5), EBS (2.0 × 2.0), EBS (2.5 × 2.5), EBS (3.0 × 3.0) and EBS (3.5 × 3.5), the mean WT was 3149ms, 3127ms, 3185ms, 3204ms and 3167ms respectively. There was also no significant interaction between character complexity and EBS ($F_{8,64} = 0.495$, $p = 0.855$).

Stroke Writing Speed (SWS)

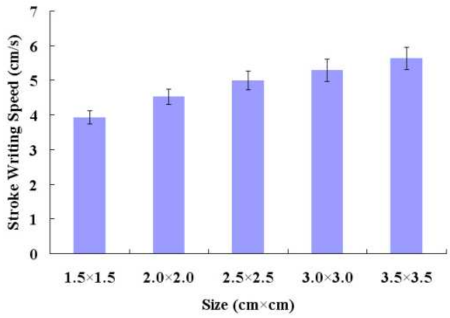


Figure 5. 9 The mean stroke writing speed for each entry box size with one-handed entry.

Repeated measures ANOVA showed that EBS had a significant main effect on SWS ($F_{4,32} = 49.531$, $p < 0.001$) (see Figure 5. 9). However, there was no interaction effect on SWS for character complexity × EBS ($F_{8,64} = 2.031$, $p = 0.057$). Post-hoc comparisons were performed using the Bonferroni adjustment for multiple comparisons, and it found no significant difference between EBS (2.5 × 2.5) (Mean = 4.997cm/s) and EBS (3.0 × 3.0) ($p = 0.046$) (Mean = 5.291cm/s). Moreover, these two EBSs were significantly faster ($p < 0.005$) than EBS (1.5 × 1.5) (Mean = 3.930cm/s) and EBS (2.0 × 2.0) (Mean = 4.530cm/s) but significantly slower than EBS (3.5× 3.5) (Mean = 5.632cm/s). Therefore, EBS (2.5 × 2.5) and EBS (3.0 × 3.0) can be grouped on the similar SWSs.

Size Ratio (SR)

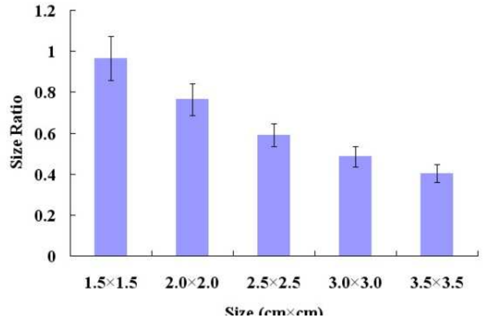


Figure 5. 10 The mean size ratio for each entry box size with one-handed entry.

Repeated measures ANOVA showed that EBS had a significant main effect on SR ($F_{4,32}$ = 41.358, $p < 0.001$). As illustrated in Figure 5. 10, smaller EBS normally led to larger SR. For EBS (1.5 × 1.5), EBS (2.0 × 2.0), EBS (2.5 × 2.5), EBS (3.0 × 3.0) and EBS (3.5 × 3.5), the mean SR was 0.964, 0.764, 0.590, 0.485 and 0.403 respectively. There was also an interaction effect on SR for character complexity × EBS ($F_{8,64}$ = 7.399, $p < 0.001$). Post-hoc comparisons were performed using the Bonferroni adjustment for multiple comparisons, and it found significant difference ($p < 0.005$) between all EBSs except EBS (1.5 × 1.5) and EBS (2.0 × 2.0) ($p = 0.127$). Therefore, EBS (1.5 × 1.5) and EBS (2.0 × 2.0) can be grouped on their similar SRs.

Length of Protruding Strokes (LPS)

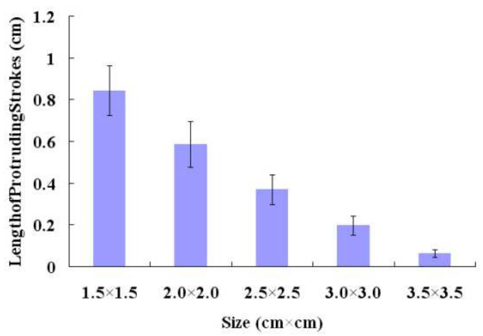


Figure 5. 11 The mean length of protruding strokes for each entry box size with one-handed entry.

An analysis of normality found that the data for LPS were skewed, so a square root transformation was used to remedy the data. Repeated measures ANOVA showed that EBS had a significant main effect on LPS ($F_{4,32} = 39.111$, $p < 0.001$). Figure 5. 11 shows that LPS tended to be reduced when EBS increased. Also, there was an interaction effect on LPS for character complexity × EBS ($F_{8,64} = 16.982$, $p < 0.001$). Post-hoc comparisons were performed using the Bonferroni adjustment for multiple comparisons. The LPS of EBS (3.5 × 3.5) was the smallest with a mean of 0.062cm, which was found to have no significant difference ($p = 0.035$) from EBS (3.0 × 3.0) (Mean = 0.197cm). There was no significant difference on LPS between EBS (3.0 × 3.0) and EBS (2.5 × 2.5) (Mean = 0.369cm) ($p = 0.010$), and also between EBS (2.5 × 2.5) and EBS (2.0 × 2.0) (Mean = 0.586cm) ($p = 0.037$). Moreover, the LPSs produced by these four EBSs are significantly smaller than the LPS produced by EBS (1.5 × 1.5) (Mean = 0.843cm), so these four EBSs can be grouped on their similar LPSs.

Number of Protruding Strokes (NPS)

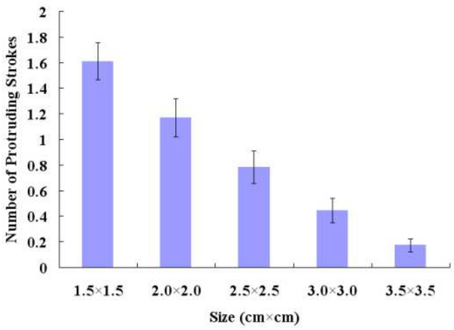


Figure 5. 12 The mean number of protruding strokes for each entry box size with one-handed entry.

Because an analysis of normality found that data of NPS were skewed, a square root transformation was applied to remedy the data. Repeated measures ANOVA showed that EBS had a significant main effect on NPS ($F_{4,32} = 110.422$, $p < 0.001$). As illustrated in Figure 5. 12, smaller EBS generally led to larger NPS. For EBS (1.5 × 1.5), EBS (2.0 × 2.0), EBS (2.5 × 2.5), EBS (3.0 × 3.0) and EBS (3.5 × 3.5), the mean NPS was 1.610, 1.169, 0.782, 0.446 and 0.172 respectively. There was also an interaction effect on NPS for character complexity × EBS ($F_{8,64} = 14.815$, $p < 0.001$). Post-hoc comparisons were performed using the Bonferroni adjustment for multiple comparisons, and it found significant difference ($p < 0.005$) between all EBSs except EBS (3.0 ×3.0) and EBS (3.5 × 3.5) (p = 0.011). Therefore, these two EBSs can be grouped with their similar NPSs.

Number of Writing Attempt (NWA)

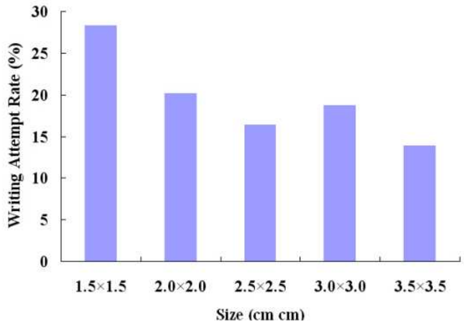


Figure 5. 13 The writing attempt rate for each entry box size with on-handed entry.

The chi-square test based on standardized residuals was applied to NWA analysis. On the frequency of NWA, no statistically significant relationship was found between NWA and EBS (χ^2_4 = 5.938, p = 0.204) (see Figure 5. 13). EBS (3.5 × 3.5) resulted in the fewest NWA with a standardized residual at -5.4. EBS (2.5 × 2.5) produced almost the same NWAs (z = -3.4) as EBS (3.5 × 3.5). EBS (2.0 × 2.0) and EBS (3.0 × 3.0) resulted in almost the same NWAs (z = 0.6 and z = -0.4 respectively). The largest NWA was produced by EBS (1.5 × 1.5) (z = 8.6).

Subjective Preferences

Dimension	EBS (1.5 × 1.5)	EBS (2.0 ×2.0)	EBS (2.5 × 2.5)	EBS (3.0 ×3.0)	EBS (3.5 × 3.5)
Writing speed	1.7	2.0	3.4	3.8	4.6
Utilization rate	4.9	4.8	4.1	3.3	2.7
Writing accuracy	1.4	2.0	3.9	4.1	4.4
Ease of writing	1.0	1.7	3.0	4.2	4.3
Fatigue of the finger	1.1	2.0	3.2	4.0	4.1

Table 5. 5 The participants' preferences for each EBS with one-handed entry.

There was a significant difference found in the effect of five EBS on the overall preference rating

($F_{4,32}$ = 47.963, p < 0.001) (see Table 5). Post-hoc comparisons using the Bonferroni adjustment showed that EBS (3.5 × 3.5), EBS (3.0 × 3.0) and EBS (2.5 × 2.5) were rated significantly higher than the other two EBSs (p < 0.005). Moreover, no significant difference was found between EBS (3.5 × 3.5) and EBS (3.0 × 3.0) (p = 1.000), EBS (3.0 × 3.0) and EBS (2.5 × 2.5) (p = 0.516), and also EBS (3.5 × 3.5) and EBS (2.5 × 2.5) (p = 0.790). Therefore, these three EBSs can be grouped with similar subjective preferences.

Learning Effect

We collected 2 blocks of data to investigate the learning effect. Repeated measures ANOVA showed no obvious learning effect on the WT ($F_{1,8}$ = 4.201, p = 0.075), SRS ($F_{1,8}$ = 2.074, p = 0.188) and LPS ($F_{1,8}$ = 4.459, p = 0.068) between the two experimental blocks. As well, no learning effect was found on the frequency of NWA (χ_1^2 = 0.01, p = 0.913) using a chi-square analysis. Although there is a significant main effect for block on NPS ($F_{1,8}$ = 9.186, p < 0.05), the overall results showed that the learning effect was minor and participants had already reached a steady performance from block one.

5.5.4 Discussion

Handwriting Speed

Writing time (WT) and Stroke writing speed (SWS) is used to quantify handwriting speed. The writing times of all EBSs were similar, but different EBSs resulted in different stroke writing speeds. Analysis results of WT and SWS suggest EBS (2.5 × 2.5) is the smallest EBS in which users can perform handwriting tasks with high handwriting speed.

Utilization Rate of Entry Area

Size ratio (SR) is applied to quantification of entry area utilization rate. The results of entry area utilization rate for one-handed entry are consistent with the results obtained in the experiment with two-handed entry; EBS (1.5 × 1.5) and EBS (2.0 × 2.0) resulted in a similar high utilization rate of the entry area. It should be noted that, the SR produced by EBS (2.5 × 2.5) was smaller than that produced by EBS (1.5 × 1.5) and EBS (2.0 × 2.0) but significantly larger than that produced by EBS (3.0 × 3.0) and EBS (3.5 × 3.5). This indicates that EBS (2.5 × 2.5) could also generate a high utilization rate of entry area.

Ease of Writing

Length and number of protruding strokes (LPS and NPS) are applied to description of ease of writing. Although the NPS produced by EBS (2.5 × 2.5) was significantly larger than that produced by EBS (3.0 × 3.0) and EBS (3.5 × 3.5), these three EBSs resulted in similar LPSs. The similar LPS was also obtained in EBS (2.0 × 2.0), but the NPS for EBS (2.0 × 2.0) was significantly larger than that for EBS (2.5 × 2.5). Therefore, EBS (2.5 × 2.5) can be regarded as the smallest entry box size in which users can perform handwriting task easily.

Entry Accuracy

The number of writing attempts (NWA) is used to describe entry accuracy. For one-handed entry, EBS (3.5 × 3.5) produced the fewest NWA. However, it should be noted that EBS (2.5 × 2.5) and EBS (3.5 × 3.5) resulted in almost the same NWA, suggesting that EBS (2.5 × 2.5) can also support accurate handwriting input.

Subjective Preferences

According to the preference ratings of these five EBSs, EBS (2.5 × 2.5) had a comparatively high overall rating. The subjective rating result was in fairly good agreement with the analysis results of handwriting speed, utilization rate of entry area, entry accuracy and ease of writing.

According to the overall analysis results of handwriting speed, utilization rate of entry area, entry accuracy, ease of writing and subjective preference, for one-handed handwriting input, the optimal entry box size was found to be 2.5cm × 2.5cm.

5.6 General Discussion and Future Work

5.6.1 Handwriting Performance within an Entry Area

Handwriting activity is both cognitive and physical, the coordinated movement of thoughts, hand and eye [94]. This study deeply explored handwriting performance within an entry area. Before designing the experiment, we expected that larger entry size would lead to greater writing speed, higher accuracy, greater ease of writing and higher subjective evaluation, but lower utilization rate of entry area; and that for smaller entry size, the result would be the opposite for all criterions. We aimed to find out the optimal entry box size in terms of a tradeoff among the above criterions. Therefore, we defined the optimal entry size for character handwriting as the smallest input area in which the user can input characters with high entry area utilization rate, great writing speed, high character recognition rates, small number and short length of stroke protrusions outside the area and high subjective assessment (for example, ease of writing and degree of fatigue). According to this definition, we defined seven dependent variables and sorted them into five measures: handwriting speed, utilization rate of entry area, accuracy, ease of writing and subjective evaluation. Inspired by previous studies ([72], [89], [79]), we performed a detailed analysis for each measure to find the optimal entry box size. For both one-handed entry and two-handed entry, the results show that there was no significant difference among EBS (2.5 × 2.5), EBS (3.0 × 3.0) and EBS (3.5 × 3.5) in terms of handwriting speed, ease of writing, entry accuracy and subjective evaluation. In addition, although utilization rate of entry area produced by EBS (2.5 × 2.5) was smaller than that produced by EBS (1.5 × 1.5) and EBS (2.0 × 2.0), it was significantly larger than that produced by EBS (3.0 × 3.0) and EBS (3.5 × 3.5), suggesting that EBS (2.5 × 2.5) could also generate a high utilization rate of entry area. Therefore, it can be concluded that EBS (2.5 × 2.5) is large enough for fast, accurate and easy handwriting with a high entry area utilization rate and subjective evaluation. This result reveals that different entry box sizes can result in different handwriting performance and highlights the importance of entry box size for handwriting input.

Handwriting speed and handwriting accuracy are two important factors in handwriting performance. High handwriting speed usually indicates that the user can write characters easily. Our results show that the user can achieve a high handwriting speed if the entry area is larger than 2.5cm × 2.5cm. Although the participants generally reported that larger entry area can lead to greater writing speed, for entry areas larger than 2.5cm × 2.5cm, there was no significant speed increase according to the experimental data analysis. Another important point in evaluation of handwriting performance is handwriting accuracy. Our results indicate that inputted characters can be recognized well if the entry area size is 2.5cm × 2.5cm or larger. The analysis of handwriting speed and length and number of protruding strokes shows that the user wrote faster with shorter and fewer protruding strokes in a large entry area than in a small one. From this, we infer that the user can write characters more clearly in a larger entry area. Therefore, inputted characters in a large area should be clear enough for the handwriting recognition software to recognize them correctly. The same can also be inferred from the subjective reports of the experimental participants; over half of them said: "The entry box with 1.5cm × 1.5cm is too small to write strokes clearly and comfortably".

5.6.2 Controlled Factors and Uncontrolled Factors in the Experiment

We carefully designed the experiment and controlled the possible confounding factors for the purpose of our study. The main controlled factors in the experiment were: 1) entry box position: the entry box was set in the center of the screen of the experimental device. According to the study conducted by Karlson et al. [38] this entry area can support fast tapping, hence may also support fast handwriting; 2) prototype Chinese characters: these characters are commonly used Chinese characters. The participants may write these characters faster and easier than they can write rarely used characters; 3) participants: the participants were younger adults and were familiar with the Chinese characters used in the experiment; these subjects may perform better than children and elder adults as well as users who are not familiar with these characters; 4) body posture: the participants were asked to sit in a chair to perform the tasks, which may lead to a better performance than standing or walking; 5) experimental device: the experiment was conducted on a HTC touch HD mini smartphone, a popular touch-based mobile phone with a HVGA screen. The conclusion that 2.5cm × 2.5cm is the optimal entry box size was drawn by using the controlled experiment setup described above. This entry box size will help the UI designer design a rational screen layout which can display more information and also allow users to write with ease and high efficiency. If the screen area is large enough, the designer can enlarge the entry box size accordingly, although this may not significantly improve the user's handwriting performance according to the results of this study. For other scenarios, such as walking, the results of this study provide important guidelines to help the UI designer set up some handwriting entry box sizes for evaluation. By examining the handwriting performance within these boxes, the designer can determine the optimal entry box size according to the methodology proposed in this study. On the other hand, the uncontrolled factors in the experiment mainly referred to the participants, such as their moods and fatigue level when performing the experiment.

Because our goal was to determine optimal entry box dimensions by measuring handwriting performance variables in different EBSs, the experiments were designed and conducted in a lab setting that allowed us to efficiently collect a large amount of data for quantitative analysis. In our experiments, the entry box position was set in the center of the screen of the experimental device. However, in mobile phones such as iPhone 4 (Apple Inc.), the entry box is set at the bottom of the screen. Nevertheless, using the methodology of our study, optimal entry box size can be determined for any entry box position. In future work, we will examine optimal entry box size in different regions of the screen. Also, we will continue to explore the impact of user age (younger or older subjects) and body posture (sitting, standing and walking) on handwriting input and we will identify the optimal entry box size for different body postures and different user age groups.

5.7 Conclusion

Two experiments were conducted to investigate the optimal finger-based entry size in touch-based mobile phones for two commonly used Chinese handwriting input styles: two-handed entry with the non-dominant hand holding the device and the index finger of the dominant hand entering characters; and one-handed entry with the dominant hand holding the device and the thumb of the dominant hand being used for character entry. A set of variables for performance measure were proposed and a detailed analysis procedure was carried out here, which enabled the determination of the optimal entry box size for handwriting input. For both one-handed entry and two-handed entry, the optimal entry box size was found to be 2.5cm×2.5cm, suggesting that this size of entry box is large enough for fast and accurate handwriting with high entry area utilization rate and few, short protruding strokes. We believe the user interface design for handwriting in touch-based mobile phones can be benefit from the experimental results and the methodology of this study.

Chapter 6 A Comparison of Flick and Ring Document Scrolling in Mobile Environments

6.1 Introduction

Advances in processing speed and memory allow mobile phones to support a number of applications such as text view and edit. However, the small screen area of mobile phones restricts the size of displayed text. Therefore, the user has to interact more fluently with the device in order to get to the desired location in the text. Thus scrolling is important for the support of many document related tasks in mobile phones.

As two commonly employed techniques for document and list navigation, flick and ring are present in a wide range of electronic devices including touch-based mobile phones and portable media players. In flick gesture a finger slides in a line along the screen. As an intuitive and natural form, flick has been commonly employed in touch-based mobile phones such as the iPhone. On the other hand, ring is used to effect document scrolling by means of circular strokes. Ring serves as an efficient scrolling technique in mobile devices such as the Apple iPod.

For efficient interaction with digital devices for scrolling documents, much research in recent years has focused on the design and analysis of flick and ring techniques [1], [21], [65], [84], [93], [109]. For example, Aliakseyeu et al. [1] systematically investigated the effectiveness of multi-flick in pen-based interfaces by designing several flick-based scrolling techniques and comparing their performance with that of a scrollbar. In the study, multi-flick technique achieved as good a performance as the scrollbar. Inspired by the hardware scrolling rings like the one in the Apple iPod, Moscovich and Hughes [65] proposed a technique for scrolling through documents by means of a virtual scroll ring.

However, to the best of our knowledge, the performances of these two techniques in touch-based mobile phones have never been directly compared in a formal evaluation. The widespread use of flick and ring for document scrolling in touch-based mobile devices signified the importance of these two scrolling techniques. An open question is which scrolling technique performs better in the context of document navigation tasks. We would like to explore the design space of scrolling techniques on mobile devices, and evaluate these designs in two mobile environments: sitting and walking. Finding the advantages and disadvantages of each scrolling technique can expedite the design of scrolling techniques and result in significant benefits to users.

This study describes two experiments in which we examined the performance of flick and ring scrolling for document navigation in touch-based mobile phones using three input methods (index finger, thumb, and pen), under two mobile environments: sitting and walking respectively. In index finger input, the non-dominant hand holds the device and the index finger of the dominant hand is used for gesturing. In thumb input, the dominant hand holds the device and the thumb of the dominant hand is used for gesturing. Although pen input is not prevalent in touch-based mobile phones, it can be an alternative to finger input in some cases such as when users operate small targets in mobile phones [77]. Hence, we also compared the performances of flick and ring with pen input. A variant of Fitts' reciprocal tapping task, which is similar to that used by Hinckley et al. [27] was used to thoroughly compare flick scrolling and ring scrolling.

This study begins with a review of related work, covering literature on flick scrolling, ring scrolling, device-independent scrolling and movement time models for scrolling. This is followed by a description of the experiment design. Then we present two quantitative experiments. After each experiment, we discuss the results for that experiment. Finally, we provide several guidelines for the design of scrolling techniques in mobile phones.

6.2 Related Work

This work builds upon four areas of previous research, most of which focused on pen-based interaction. The first refers to the flick scrolling technique. The second is a body of work on ring scrolling technique. The third one is about device-independent scrolling. The last is movement time models for scrolling. We review each in turn.

6.2.1 Flick Scrolling Technique

Flick scrolling is an intuitive and natural scrolling method for mobile touch devices such as the iPhone. Aliakseyeu et al. [1] designed four flick-based scrolling (multi-flick) techniques and compared them with the traditional scrollbar for navigating lists and documents on different devices (PDA, tablet PC, large table). In the study, multi-flick technique achieved as good performance as the scrollbar. Yin and Ren [114] used pen pressure to improve the performance of flick-based and ring-based scrolling techniques. Experimental results indicated that the techniques with pressure information performed better than that without pressure information. However, they did not pay attention to the comparison of flick and ring techniques.

6.2.2 Ring Scrolling Technique

Ring gesture is a circular motion. Rotating scroll wheels is one of the most widely adopted scrolling techniques in devices such as iPod. Earlier work by Wherry [109] investigated the performance of a touchpad scroll ring, a mouse scroll wheel and touchpad scroll zone in a variant of Fitts' tapping task; the scroll ring performed faster with fewer errors. To improve list selection performance, Diehl et al. [21] designed a novel scroll ring with pressure sensitivity. Results indicated this scrolling ring could offer a potential interaction advantage.

Inspired by the hardware scrolling ring such as that in the Apple iPod, Moscovich and Hughes [65] proposed a technique for scrolling through documents by means of a virtual scroll ring. The technique used the amplitude and frequency of repetitive circular movement, rather than the angle and radius, to better support ring document scrolling. Results showed that VSR performed at least as well as a mouse wheel for medium and long distances, and was preferred by users.

In order to better support scrolling on touch displays, Smith and Schraefel [93] designed a radial scroll widget: the scrolling time for the scroll widget was shorter than that for the traditional scrollbar for short scrolling distance. However, the scroll widget suffered a drawback that the user must maintain visual focus on it. Curve dial [84] can support eyes-free parameter entry for document scrolling, as it tracked the curvature arc rather than the center. Radial scroll tool and curve dial selected a minimum of three points to determine the angle of curvature, which inspired the ring technique design in our study.

6.2.3 Device-independent Scrolling

As an easy to implement technique, the scrollbar has been widely used for navigating documents in a wide range of electronic devices including computers, graphing calculators, mobile phones, and portable media players. Igrashi and Hinkley [35] proposed a novel navigation technique for browsing large documents, speed-dependent automatic zooming (SDAZ), which use scrolling combined with an automatic zooming mechanism to provide fast visual search. A user study demonstrated the effectiveness of their technique. Flipper [95] is a variation of SDAZ technique, which can enable the user to scroll at high rates one page at a time. Cockburn et al. [19] proposed displacement-dependent automatic zooming (DDAZ), which can be used to navigate documents by scrolling and zooming in proportion to the amount of cursor displacement. Results showed that DDAZ was faster for scrolling than SDAZ. From a view of eliminating most scrolling, Cockburn et al. [18] proposed Space-Filling Thumbnails (SFT), which aims to provide fast document navigation by using an overview display.

6.2.4 Movement Time Model for Scrolling

A quantitative human performance model would facilitate the design and evaluation of scrolling techniques by quantitatively predicting their efficiency before running extensive user studies. In a early study, Zhai et al. [117] investigated the performance of three input methods (mouse with isometric joystick, mouse with a track wheel, and two handed joystick and mouse) in a task that involved both scrolling and pointing. The results showed that a mouse with a finger wheel did not improve user’s performance but the other input methods significantly improved users' performance. In a noteworthy analytical study, Hinckley et al. [27] have shown that Fitts' law can model certain scrolling patterns. In the study, participants were asked to perform a variant of Fitts' reciprocal tapping task by means of a IBM ScrollPoint and a IntelliMouse Wheel. However, the study did not examine the applicability of Fitts' law for ring and flick document scrolling in touch-based mobile phones. Another movement time model for scrolling was proposed by Andersen [2] (Andersen's model), taking into account that Fitts' law was developed for "aimed" movement but for scrolling tasks the target position is usually not known. The study indicated that movement time was linearly dependent on the target distance. In our study, we further examined the effectiveness of Fitts' Law and Andersen model for the prediction of movement time with ring and flick document scrolling in touch-based mobile phones.

6.3 Method

6.3.1 Flick and Ring Techniques

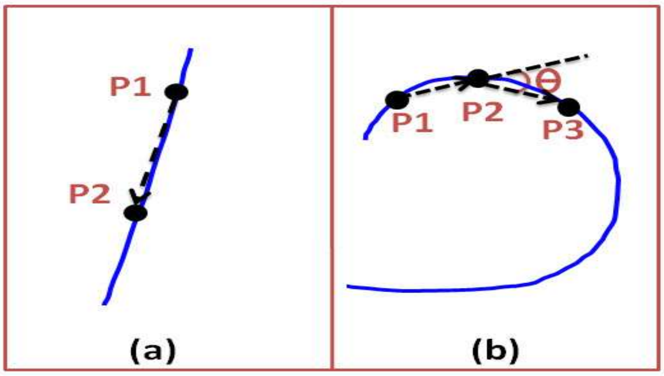


Figure 6. 1 The illustration of (a) flick scrolling and (b) ring scrolling.

The flick technique used here was designed based on the method proposed in [114]. As illustrated in Figure 6. 1a, p2(x2, y2) and p1(x1, y1) respectively denote the current and previous points in a gesture trajectory. The document scrolling distance is equal to the absolute value of (y2 - y1). The document scrolling direction is determined by the sign of (y2 - y1): if the sign is negative, the document will scroll forward. Otherwise, the document will scroll backward.

On the other hand, for ring technique, we utilized a method similar to that used in [84], [93], [114]. As illustrated in Figure 6. 1b, there are a minimum of three points p1, p2 and p3 (p1 is a previous point of p2, and p2 is a previous point of p3) in a gesture trajectory. θ denotes the angle that rotates from the vector (p1, p2) to the vector (p2, p3). The document scrolling distance is equal to $\theta \times R/2\pi$ (R is a constant with a value of 220 pixels). The scrolling direction is determined by the sign of the dot product of the vector (p1, p2) and the vector (p2, p3): if the sign is positive, the document would scroll forward. Otherwise, the document will scroll backward. Scrolling by angle indicates that fast and small circles can cause fast scrolling, while slow and large circles can cause slow scrolling.

6.3.2 Reciprocal Framing Task for Scrolling

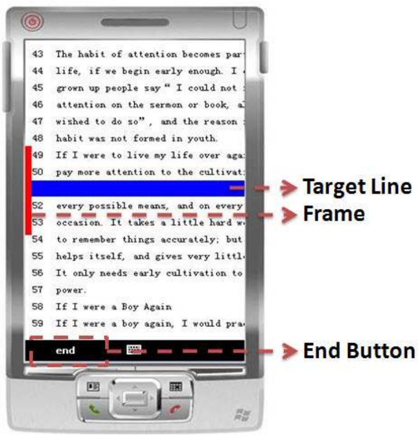


Figure 6. 2 Experimental interface.

The experimental task was similar to [27], which was a variant of Fitts' reciprocal tapping task. In the experiment, participants were instructed to scroll down and up, moving back and forth between two target lines in a document using flick or ring technique. As illustrated in Figure 6. 2, a document consisted of 288 lines with a line height of 21 pixels (0.30 cm) was used. We assigned every line a unique number, starting at 1 for the first line and incrementing by 1 for each successive line. We expected that these numbers can help participants to find the target lines easily. The initial target line appeared with red and the second one was marked by blue. A frame was placed at the left of the task window and always centered on the screen. Participants were asked to scroll a target line toward the range of the screen range identified by the frame. Once the target line fully entered the identified screen range, participants were asked to press the end button in order to complete the current scrolling and continue the next scrolling; meanwhile, the target line will disappear. If participants pressed the end button without the target line fully entering the identified screen range, a warning beep tone would be present to them, but we asked them to continue to scroll toward the next target.

6.4 Experiment One: in Sitting Posture

6.4.1 Participants

Ten participants, 9 males and 1 female, from 20 to 27 years of age, took part in this experiment. All of them were right-handed and had prior experience with bare finger operation on touch screen devices such as iPhone. Six of them had prior experience operating digital screens with digital styli.

6.4.2 Apparatus

The study was conducted on a HTC Touch HD mini smart phone equipped a capacitive touch screen with HVGA resolution. The screen size is 3.2 inch and its resolution was 320 × 480 pixels. The platform is Windows Mobile 6.5 Professional with HTC Sense. The experimental program was designed in the C# environment.

6.4.3 Task and Procedure

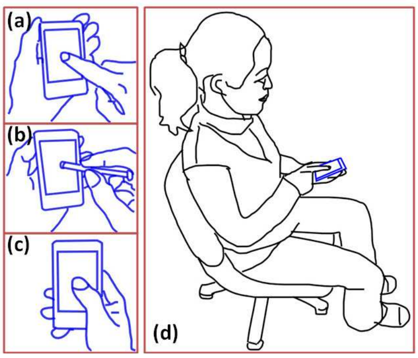


Figure 6. 3 (a) Index finger input. (b) Pen input. (c) Thumb input. (d)Participant in the experimental environment

When performing the experimental task, participants were asked to sit in a chair (see Figure 6. 3d). The experiment used a 3 × 2 × 4 × 3 within-factor design with a variety of planned comparisons. The independent variables were input method (index finger, pen and thumb, see Figure 6. 3a, b, c respectively), scrolling technique (flick and ring), target distance (20, 60, 120 and 200 lines), and frame width (3, 6 and 12 lines). A (partially-balanced) Latin-square was used to counterbalance the order of the presentation of the input method and scrolling technique. For each input method and scrolling technique, the order of the 4 target distances for the 3 frame widths was randomized. For each target distance and frame width, the participants completed 7 individual target acquisitions (phase). The participants on average took 60 minutes to complete the experiment. In summary, experiment data collection consisted of:

10 subjects ×

3 input methods ×

2 scrolling techniques ×

4 target distances ×

3 frame widths ×

7 phases

= 5040 scrolling trials

At the end of the experiment, a questionnaire was administrated to gather subjective opinions. Participants were asked to rate flick and ring for each input method on 7-point Likert Scales regarding *movement speed*, *easy to position target line* and *hand fatigue* (7 for highest preference, and 1 for lowest preference).

6.4.4 Results

Learning Effects

Recall that for each target distance and frame width, each participant performed 7 phases. In order to ensure data stability, we first checked the learning effect on movement time over the 7 phases to see if the data we collected had reached a level of stability. Movement time is defined as the duration from the moment the pen or finger touches the screen to the moment the target line last enter the region specified by the frame before the “end” button is pressed.

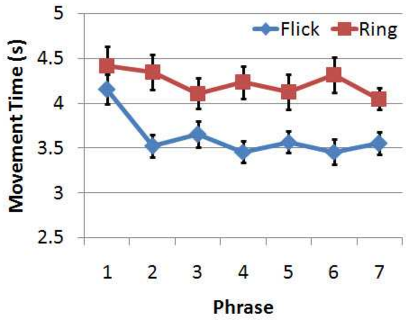


Figure 6. 4 Mean movement time for each phase and scrolling technique.

As shown in Figure 6. 4, for flick technique, repeated measures ANOVA showed that phase had a significant main effect on movement time ($F_{6, 54}$ = 12.43, p < 0.001). Post-hoc comparisons revealed that the first phase had a significant longer movement time than the other phases (p < 0.05). Therefore, we excluded the data the first phase for the rest of our analysis. On the other hand, in respect to ring technique, although no significant main effect was found on movement time for phase ($F_{6, 54}$ = 2.06, p = 0.07), it was found the first phase resulted in a significant longer movement time than the third, fifth and seventh phase (p < 0.05 for all); in these four phases, the experimental tasks were the same: moving the red target line into the region specified by the frame. Hence, data excluded the first phase was applied to the rest of our analysis.

Number of Crossings (NC)

When moving the target line into the screen region specified by the frame, participants sometimes crossed the frame more than once. Number of crossings is defined as the number of times the target line enters or leaves the specified frame region for a particular trail with one target distance and frame width, minus 1.

Regarding index finger input, repeated measures ANOVA showed a significant main effect on NC for scrolling technique ($F_{1, 9}$ = 107.37, p < 0.001). The mean NC was 2.14 in the flick condition and 5.05 in the ring condition for each target distance and frame width. Other independent variables influenced NC. A significant main effect was found on NC for target distance ($F_{3, 27}$ = 7.08, $p < 0.01$) and frame width ($F_{2, 18}$ = 47.49, $p < 0.001$). Interesting, although there was no significant interaction effect on NC for frame width, there was an interaction between scrolling technique and target distance ($F_{3, 27}$ = 3.76, $p < 0.05$).

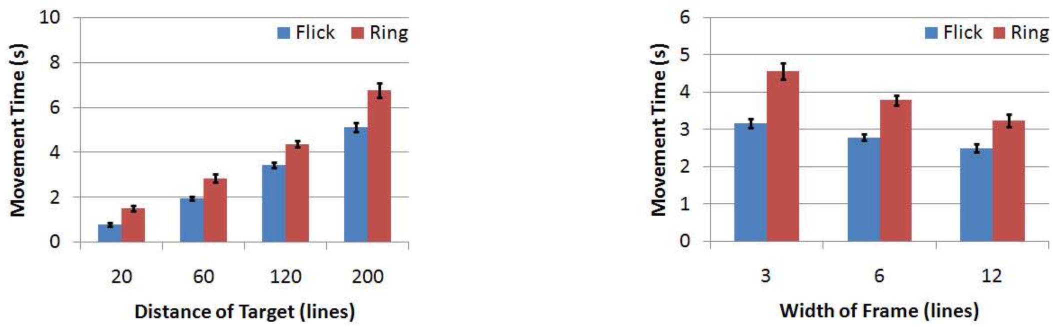
For pen input, there was a significant main effect on NC for scrolling technique ($F_{1, 9}$ = 41.90, $p < 0.001$). The mean NC was 1.83 in the flick condition and 5.04 in the ring condition for each target distance and frame width. There was a significant main effect on NC for target distance ($F_{3, 27}$ = 5.88, $p < 0.01$) and frame width ($F_{2, 18}$ = 47.49, $p < 0.001$). Although no significant interaction effect was found between scrolling technique and target distance, there was a strong interaction between scrolling technique and frame width ($F_{2, 18}$ = 4.14, $p < 0.05$).

With respect to thumb input, a significant main effect was found on NC for scrolling technique ($F_{1, 9}$ = 75.21, $p < 0.001$). The mean NC was 0.73 in the flick condition and 3.08 in the ring condition for each target distance and frame width. The frame width had a significant main effect on NC ($F_{2, 18}$ = 35.07, $p < 0.001$). Although there was no significant interaction between scrolling technique and target distance, there was a strong interaction between scrolling technique and frame width ($F_{2, 18}$ = 8.66, $p < 0.05$).

Overall, for index finger, pen and thumb input, ring technique resulted in more NC than flick technique, indicating it is difficult for participants to position the target line within the frame using ring technique. More NC indicated that it would take longer time to position the target line within the frame, which may yield a false measurement of the movement time of scrolling. Therefore, to analyze movement time, we excluded the data of the experimental trial in which NC was greater than 3.

Movement Time (MT)

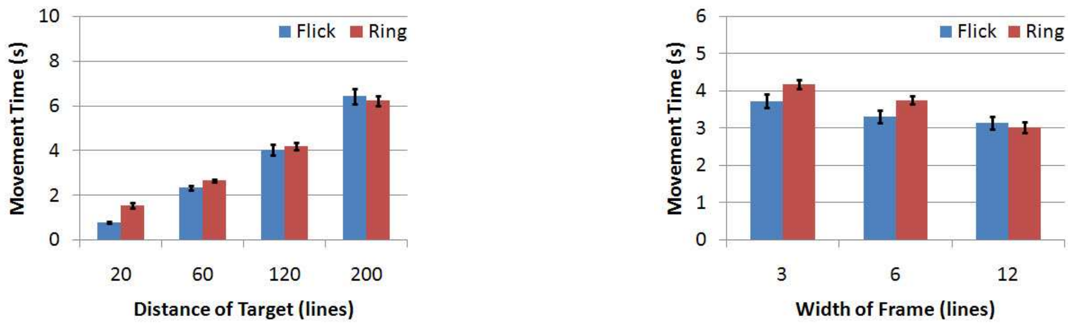
Movement time, as defined in the section “learning effects”, is another basic measure of the performance of scrolling technique.



(a) (b)

Figure 6. 5 Regarding index finger input, mean movement time for two scrolling techniques for each (a) target distance and (b) frame width. Error bars represent 0.95 confidence interval.

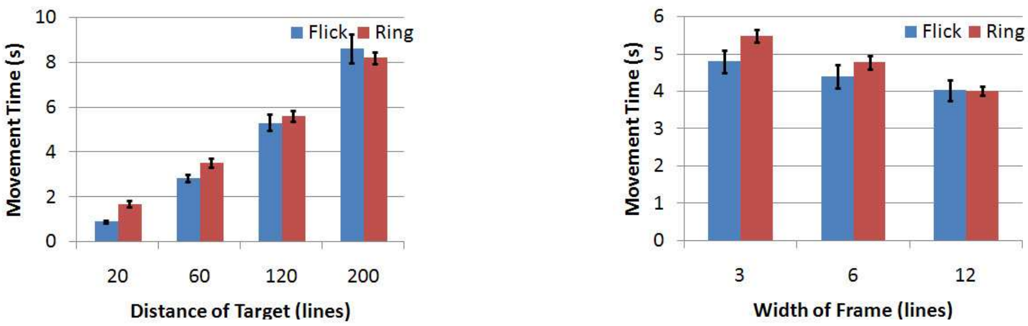
For index finger input, repeated measures ANOVA showed that scrolling technique had a significant main effect on MT ($F_{1, 9}$ = 51.54, p < 0.001). The mean MT was 2.82s for flick technique and 3.86s for ring technique. Other independent variables influenced MT. A significant main effect was found on MT for target distance ($F_{3, 27}$ = 567.55, p < 0.001) and frame width ($F_{2, 18}$ = 87.23, p < 0.001). There was an interaction between scrolling technique and target distance ($F_{3, 27}$ = 6.65, p < 0.01) (see Figure 6. 5a). In addition, there was a significant interaction between scrolling technique and frame width ($F_{2, 18}$ = 14.59, p < 0.001) (see Figure 6. 5b). The results indicated that flick performed faster than ring when users performed scrolling task by means of the index finger.



(a) (b)

Figure 6. 6 Regarding pen input, mean movement time for two scrolling techniques for each (a) target distance and (b) frame width.

In respect to pen input, there was no significant main effect on MT for scrolling technique ($F_{1, 9}$ = 2.59, p = 0.14). The mean MT was 3.39s for flick technique and 3.65s for ring technique. There was a significant main effect on MT for target distance ($F_{3, 27}$ = 588.87, p < 0.001) and frame width ($F_{2, 18}$ = 98.64, p < 0.001). A significant interaction effect on movement time was found between scrolling technique and frame width ($F_{2, 18}$ = 12.12, p < 0.001) (see Figure 6. 6b). Interesting, as illustrated in Figure 6. 6a, there was also an interaction between scrolling technique and target distance ($F_{3, 27}$ = 3.85, p < 0.05). Ring resulted in a longer MT than flick for the target distance with 20, 60, 120 lines, but shorter MT for the target distance with 200 lines.

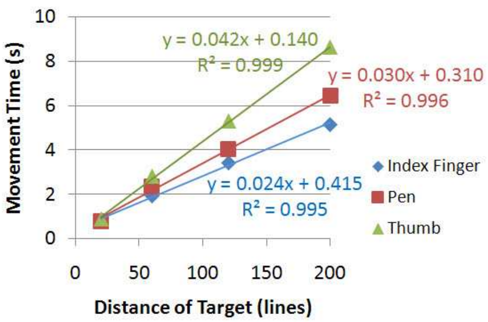


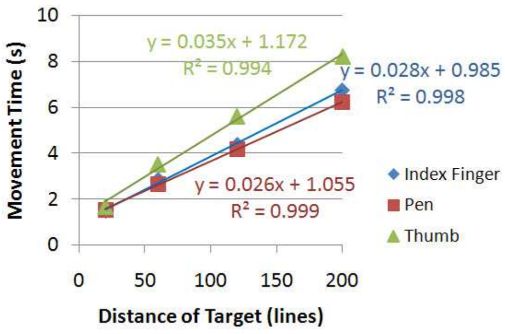
(a) (b)

Figure 6. 7 Regarding thumb input, mean movement time for two scrolling techniques for each (a) target distance and (b) frame width.

With regard to thumb input, there was no significant main effect on MT for scrolling technique ($F_{1, 9}$ = 0.92, p = 0.36). The mean MT was 4.40s in the flick condition and 4.75s in the ring condition. Other independent variables influenced MT. It was found that there was a significant main effect on MT for target distance ($F_{3, 27}$ = 383.83, p < 0.001) and frame width ($F_{2, 18}$ = 71.27, p < 0.001). There was a significant interaction between scrolling technique and frame width ($F_{2, 18}$ = 6.69, p < 0.01) (see Figure 6. 7b). Additionally, an interaction effect on movement time was found between scrolling technique and target distance ($F_{3, 27}$ = 3.11, p < 0.05) (see Figure 6. 7a). Ring produced a longer MT than flick for the target distance with 20, 60, 120 lines, but shorter MT for the target distance with 200 lines.

The Fit of Fitts' Law and Andersen Model





(a) (b)

Figure 6. 8 Movement time for each target distance in (a) flick scrolling and (b) ring scrolling.

We examined the relationship between movement time and scrolling task with respect to each scrolling technique and each input method. It was found that for each scrolling technique and input method, linear regression of the movement time by ID showed low correlations with Fitts' law ($R^2 < 0.8$). However, as shown in Figure 6. 8, for Andersen model [2], regression of D against movement time for each scrolling technique and input method yielded a good fit, with regression coefficients of 0.99. The results verified the applicability of Andersen model in flick-based and ring-based scrolling in touch-based mobile phones.

Error Rates

The error rate was defined as the percentage of target acquisition trials in which the participants pressed the “end” button but the target line was not in the range specified by the frame. It was found that for each input method and scrolling technique, the error rate was very low with a value less than 4%.

Subjective Evaluation

Repeated measure ANOVA showed that for index finger input, flick was rated significantly higher than ring in terms of movement speed, easy to position target line and hand fatigue ($p < 0.05$

for all). In respect to pen input and thumb input, although there was no main effect on *movement speed* and *hand fatigue* for scrolling technique, scrolling technique had a significant effect on *easy to position target line* ($p < 0.05$); flick was rated higher than ring. The subjective preference was fairly consistent with the movement time and number of crossings performance.

6.5 Experiment Two: in Walking Posture

Users sometimes rely on mobile devices to view and edit documents while walking. Hence, it is important to examine the performance of flick and ring in walking environments as well.

6.5.1 Participants and Apparatus

The same ten subjects who participated in experiment one took part in experiment two. The same mobile phone was used as in experiment one. We asked the participants to walk on a treadmill when performing the experiment task.

6.5.2 Task and Procedure

A similar task and procedure used in Experiment one was carried out in this experiment. When performing the experimental task, participants were asked to walk on the treadmill (walking speed was set as 0.91 m/s according to [42]). The experiment used a 3 × 2 × 3 × 3 within-factor design with a variety of planned comparisons. The independent variables were input method (index finger, pen and thumb), scrolling technique (flick and ring), target distances (20, 60 and 200 lines), and frame widths (3, 6 and 12 lines). A (partially-balanced) Latin-square was used to counterbalance the order of the presentation of the input method and scrolling technique. For each input method and scrolling technique, the order of the 3 target distances for the 3 frame widths was randomized. For each target distance and frame width, the participants completed 4 individual target acquisitions (phase). The participants on average took 30 minutes to complete the experiment. In summary, experiment data collection consisted of:

10 subjects ×

3 input methods ×

2 scrolling techniques ×

3 target distances ×

3 frame widths ×

4 phases

= 2160 scrolling trials

It should be noted that, in this experiment, in order to avoid user fatigue after a long walk, we only selected 3 target distances, rather than the 4 used in Experiment one. Also the number of phases was reduced to 4. At the end of the experiment, a questionnaire was administered to gather subjective opinions. Participants were asked to rate flick and ring for each input method on 7-point Likert Scales regarding *movement speed, easy to position target line* and *hand fatigue* (7 for highest preference, and 1 for lowest preference).

6.5.3 Results

Learning Effects

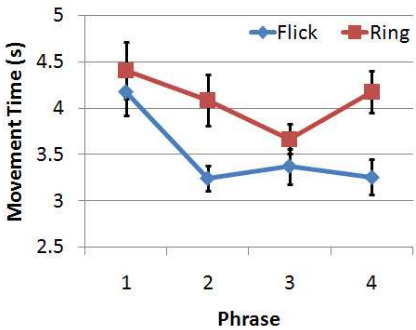


Figure 6. 9 Mean movement time for each phase and scrolling technique.

Subjects performed the same process as in Experiment one and we first examined the learning effects for further data analysis. Figure 6. 9 illustrates average movement time for each phase regarding each scrolling technique. For both flick and ring technique, a repeated measures ANOVA analysis showed that phase had a significant main effect on movement time ($p < 0.001$). With respect to flick, post-hoc comparisons revealed that the first phase had a significantly longer movement time than the other phases ($p < 0.01$). Hence the data in the first phase was excluded for the rest of our analysis. With respect to ring technique, it was found that the first phase resulted in a significantly longer movement time than the third phase ($p < 0.05$ for all). Moreover, in the two phases, the experimental tasks were the same; moving the red target line into the region specified by the frame. As a result, data excluding the first phase was applied to the rest of our analysis.

Number of Crossings (NC)

Regarding index finger input, a repeated measures ANOVA analysis showed a significant main effect on NC for scrolling technique ($F_{1, 9} = 32.61$, $p < 0.001$). The mean NC was 2.12 in the flick condition and 5.23 in the ring condition for each target distance and frame width. Other independent variables influenced NC. A significant main effect was found on NC for target distance ($F_{2, 18} = 10.91$, $p < 0.01$) and frame width ($F_{2, 18}$ = = 30.04, $p < 0.001$).

For pen input, there was a significant main effect on NC for scrolling technique ($F_{1, 9} = 103.12$, $p < 0.001$). The mean NC was 2.00 in the flick condition and 5.17 in the ring condition for each target distance and frame width. There was a significant main effect on NC for target distance ($F_{2, 18} = 7.82$, $p < 0.01$) and frame width ($F_{2, 18} = 100.24$, $p < 0.001$).

With respect to thumb input, a significant main effect was found on NC for scrolling technique ($F_{1, 9} = 16.40$, $p < 0.01$). The mean NC was 0.73 in the flick condition and 3.31 in the ring condition for each target distance and frame width. The frame width had a significant main effect on NC ($F_{2, 18} = 20.79$, $p < 0.001$).

For all three input methods, no interaction effect was found on MT for target distance or for frame width.

In summary, ring technique led to more NC than flick technique for index finger, pen and thumb input. For the analysis of movement time, we excluded the data of the experimental trial in which NC was greater than 3.

Movement Time (MT)

For index finger input, repeated measures ANOVA showed that scrolling technique had a significant main effect on MT ($F_{1,\,9}$ = 7.02, $p < 0.05$) (see Figure 6. 10). The mean MT was 2.85s for flick technique and 4.04s for ring technique. Other independent variables influenced MT. A significant main effect was found on MT for target distance ($F_{2,\,18}$ = 102.22, $p < 0.001$) and frame width ($F_{2,\,18}$ = 6.69, $p < 0.01$). The results indicated that flick performed faster than ring when users performed scrolling tasks by means of index finger.

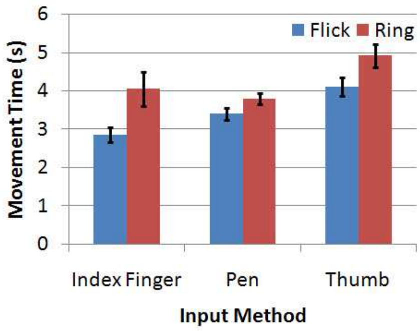


Figure 6. 10 Mean movement time for two scrolling techniques (flick and ring) for each input method (index finger, pen and thumb)

In respect to pen input, there was no significant main effect on MT for scrolling technique ($F_{1,\,9}$ = 3.10, p = 0.11) (see Figure 6. 10). The mean MT was 3.99s for flick technique and 4.29s for ring technique. There was a significant main effect on MT for target distance ($F_{2,\,18}$ = 352.82, $p < 0.001$) and frame width ($F_{2,\,18}$ = 32.08, $p < 0.001$).

With regard to thumb input, there was a significant main effect on MT for scrolling technique ($F_{1,\,9}$= 6.55, $p < 0.05$) (see Figure 6. 10). The mean MT was 4.11s in the flick condition and 4.91s in the ring condition. Other independent variables influenced MT. It was found that there was a significant main effect on MT for target distance ($F_{2,\,18}$ = 186.84, $p < 0.001$) and frame width ($F_{2,\,18}$ = 8.40, $p < 0.01$).

For index finger, pen and thumb input, there was no interaction effect on MT for target distance or for frame width.

In summary, flick achieved significantly shorter movement times than ring with index finger input and thumb input.

The Fit of Fitts' Law and Andersen Model

In the walking environment, we examined the relationship between movement time and scrolling task with respect to each scrolling technique and each input method. It was found that for each scrolling technique and input method, linear regression of the movement time by ID showed low correlations with Fitts' law ($R^2 < 0.8$). However, for Andersen’s model [2], regression of D against movement time for each scrolling technique and input method yielded a good fit, with regression coefficients of 0.99. The results verified the applicability of Andersen model in flick-based and ring-based scrolling in touch-based mobile phones while walking.

Error Rates

Flick and ring led to a very low error rate in the context of document navigation by means of pen, index finger and thumb input (less than 4%).

Subjective Evaluation

A repeated measure ANOVA analysis showed that for pen, thumb and index finger input, flick was rated significantly higher than ring in terms of movement speed, ease of target positioning and hand fatigue ($p < 0.05$ for all).

6.6 Discussion

6.6.1 Advantages and Disadvantages of Flick and Ring

Flick and ring scrolling techniques were examined in the context of mobile document navigation tasks in sitting and walking activities respectively.

	Pen	Index Finger	Thumb
Sitting	No	Yes	No
Walking	No	Yes	Yes

Figure 6. 11 A summary of movement time analysis of flick and ring for each input method (pen, index finger and thumb) and for each posture (sitting and walking). “No” represents that there is no significant difference between ring and flick regarding movement time, while “Yes” represents that there is a significant difference.

As illustrated in Figure 6. 11, in sitting environment, for index finger input, flick resulted in shorter movement time than ring and was preferred by the participants, indicating that flick is a superior technique for document scrolling. Flick gesture is analogous to a throwing motion in the real world while ring gesture is a circular motion. Therefore, compared to ring gesture, flick gesture may be more natural and intuitive for index finger input. Regarding pen and thumb input, no significant difference was found on movement time for ring and flick techniques. The interaction effects on movement time for scrolling technique may have design implications. Flick led to shorter movement time than ring for short scrolling distances (target distance $\leq$ 120 lines). However, when scrolling actions were longer than 200 lines, ring tended to be faster than flick. With respect to pen input, the greater degrees of freedom afforded by pen may allow participants to use ring technique more comfortably than flick technique. For thumb input, the thumb's movement range on the screen may have an effect on scrolling performance; it is difficult to move the thumb up and down but easy to rotate it along a circle [29]. Overall these interaction effects suggest that ring is a promising scrolling mechanism for pen and thumb input.

In the walking environment, as shown in Figure 6. 11, flick produced significant shorter movement times than ring with index finger and thumb input. In addition, in the process of the experiment, 8/10 participants stated that it was more difficult to use ring than to use flick when walking, because they felt that moving the finger up and down was easier than moving the finger around a circle on the screen. The experimental analysis and the subject evaluation indicated that flick can serve as an effective and a preferred scrolling technique for users when walking.

Overall, the results above reveal that flick differed from ring in the context of document navigation tasks by means of pen input and finger input. Future scrolling technique design should exploit the advantages and avoid the disadvantages of ring and flick. Also, our study provides a methodology to help the designer better examine the performance of scrolling techniques for scrolling technique design.

6.5.2 Smooth Scrolling for Ring Technique

Smooth scrolling is a feature used to reduce what the user would perceive as “jumps” (discontinuous movement) of a document. However, in this study, number of crossings (NC) was larger for ring than for flick technique, indicating that the participants could not perform ring smoothly. Also, participants reported that it was more difficult with ring than with flick to position the target line within the region specified by the frame. As expected, larger NC led to longer movement time. Therefore, for more effective ring scrolling technique design, it is better to increase the smoothness of response to sample points. It should be noted that for the data analysis of movement time, the data of the experimental trial in which NC was greater than 3 were excluded. After doing this, we believe that even after the increase of the smoothness, the experimental results would not change.

As introduced in the Section 6.3.1, the document scrolling distance was calculated according to the angle between two vectors indicating current and previous finger or pen positions on the screen. With respect to the ring scrolling mechanism, sufficiently fine control of ring was difficult and as a result, participants could not achieve smooth scrolling. Several methods have been proposed to support smooth scrolling, including linear least-squares fit [65], and increasing the gap between sampling points selection [84]. These methods will be used in our future examination of the performance of ring scrolling.

6.5.3 Mapping Function for Flick and Ring techniques

Aliakseyeu et al. [1] proposed three mapping functions for flick technique and demonstrated their effectiveness in the context of document and list navigation tasks. However, we did not use those mapping functions for two reasons. First, we wanted to avoid complicating our results with different varieties of mapping functions in the preliminary investigation of the performance of flick and ring, so we designed our scrolling techniques based on a simple mapping function which performed a linear translation of the displacement of input method to the distance of document scrolling. Second, the aim of this study was to compare flick and ring scrolling techniques, so it is essential to use a mapping function which is “fair” for both scrolling techniques. As the mapping functions in [1] were designed for evaluating flick scrolling only, they may favor flick over ring. Instead, we used the linear mapping function proposed in [84], [93], [114] to provide a fundamental mechanism which was fair for both ring and flick. Future work should further explore the performance of flick and ring document scrolling in context of different mapping functions. As a fundamental study, our study provides a methodology and some important conclusions which can be beneficial to the further investigation.

For ring technique, *R*, a constant coefficient, plays an important role in determining scrolling speed. Hence, we conducted a pilot study to select a proper *R*. We designed three ring techniques, in which *R* was set as 110, 220 and 330 pixels respectively (R110, R220 and R330 were used to denote the three ring techniques respectively). Six participants were asked to perform these three ring techniques. The experiment procedure was similar to that introduced in section “experiment design”. As a result, R220 resulted in significant shorter movement time than R110 but fewer NC than R330. In addition, no significant difference was found on movement time between R330 and R220. Hence, *R* was set as 220 pixels for our study.

6.7 Conclusion

A controlled experiment was presented here, which empirically evaluated the performance of two commonly used scrolling techniques (flick and ring) for document navigation by means of index finger input, pen input and thumb input in touch-based mobile phones, with regard to users' sitting and walking activities respectively. We found that flick performed better than ring for the three input methods. Also, with regard to pen input and thumb input, ring performed faster than flick for long target distance, indicating ring has a potential interaction advantage and should be deeply explored for future scrolling technique design. Additionally, both flick and ring document scrolling in touch-based mobile phones can be modeled by Anderson model [2] in both sitting and walking activities. These findings may be useful in improving the performance of flick and ring document scrolling in touch-based mobile phones.

Chapter 7 Window Avatar: Leveraging Interactions on Coordination in Multi-touch Tabletop Displays

7.1 Introduction

Multi-touch tabletops are becoming more widely used and appealing in recent years. Especially, large format multi-touch displays are useful for co-located collaborative work in which two or more users interact simultaneously with a shared display. Hence, they have been widely employed to support a variety of collaborative activities, such as planning, scheduling, brainstorming, design, and layout activities.

Unfortunately, current multi-touch tabletops suffer from some drawbacks in interactive manipulation and collaboration process. First, most tabletops only provide users with a single shared workspace which can not be divided or reconfigured. Therefore, it is difficult to establish a territory which can support sharing of documents and collaboration while maintaining user control over documents and gradations of privacy [87]. Second, users in different sides of a tabletop have different views of the display, thus can not share the same orientation, making it difficult to comprehend objects, collaboration and communicate [47]. Third, accessing remote territories that are out of reach requires that the user has to walk or to use supplemental techniques, which produces a slow and interrupted workflow [10].

A number of techniques have been proposed to improve the interaction on manipulation and collaboration with the consideration of the drawbacks above. For example, Scott et al. [86] proposed Storage Bins for dealing with workspace clutter. Shen et al. [91] proposed rotatable widgets which support reorientation to different positions. Bragdon et al. [10] employed gesture to help users acquire remote targets. These techniques have been shown to be effective; however, they each address only one part of the overall problem faced by users of tabletop systems. In practice, users may have a high working load of changing different techniques. Furthermore, these techniques were mainly proposed to meet the requirements of manipulation and collaboration on pen-based tabletop displays. For multi-touch tabletop displays, which can support more natural and direct interaction by allowing fingers from both hands to be used together, these techniques may be not suitable for manipulative and collaborative interactions. Hence, it is important to design efficient methods which can better support interactions on multi-touch tabletop displays with the consideration of the use of natural and direct interaction properties of multi-touch input.

Inspired by the previous studies [7], [74], [86], we proposed a technique, Window Avatar, which allows the user to create a personal territory to enhance interactions on manipulation and collaboration in multi-touch tabletops (see Figure 7. 1). Current existing techniques avoid the use of WIMP paradigm for multi-touch tabletops. However, we explored the utility of window-based interactions for multi-touch interactions. Based on Window Avatar, we presented a set of interactions by means of a set of gestures. Users can employ these interactions to augment existing current collaborative manipulations, thus providing users with reconfigured workspace, easy remote territory access and gradations of privacy.

7.2 Related Work

This work builds upon two distinct areas of previous research. One refers to shape-based gestures on touch-based tabletops, the other is a body of work on territoriality in collaborative tabletops.

7.2.1 Hand Shape Based Gesture

Hand shape based gestures have been widely employed to enhance interactions in multi-touch tabletops. Previous studies mainly focused on the exploration of the use of hand gesture, and also the design of help systems for user learning of hand gesture. Regarding the exploration of the use of hand gesture, Rekimoto [76] described hand-shape based manipulations on the basis of a new sensor architecture which was also proposed by him for making interactive surfaces. Cao et al. [13] proposed ShapeTouch, in which they explored the usability of contact shape on interactive surfaces for object manipulation. Wu and Balakrishnan [112] proposed a variety of multi-finger and hand gesture based interaction techniques for multi-touch displays, and demonstrated these techniques with a prototype room furniture layout application. Based on the detection of three shape-based gestures, Rock & Rails technique was proposed to augment existing direct touch manipulation. On the other hand, for the design of help systems for user learning with hand gesture, Freeman [23] presented ShadowGuides, a system for learning of multi-touch and whole-hand gestures on multi-touch surfaces. Our study first employed shape-based gesture in combination with direct-manipulations to help users better perform manipulative and collaborative activities in multi-touch tabletops.

7.2.2 Territoriality in Collaborative Tabletops

In a noteworthy analytical study, Scott et al. [87] systematically investigated the territoriality in collaborative tabletop workspaces. They found that collaborators use three types of tabletop territories to support their interactions: personal, group and storage territories; and partitioning territory plays an important role in user collaboration. Tse et al. [99] investigated how users employ spatial separation and partitioning to avoid interference in signal display groupware. Motivated by the work practice of territoriality, several studies were conducted which aim to improve interactions on tabletops. For example, Scott et al. [86] proposed Storage Bins, a mobile storage for collaborative tabletop displays. By means of TableTrays, the user can divide the work area into visually distinct regions, so as to better manage space, object, and collaboration in tabletops [74]. Klinkhammer et al. [41] designed a tracking system for tabletops, which can provide data on a user's location and movement. Based on the system, an adaptive personal territory on the tabletop can be provided to the user. Tse et al. [99] explored the design space of a split view tabletop for three types of existing applications: independent applications, shared screen and true groupware. Inspired by the metaphor of “piles”, Bauer et al. [7] presented a self-adjusting clump and a grid-layout tray for managing digital photo collections.

Our review showed that no study paid attention to augment collaborative interactions on multi-touch tabletops involving the consideration of partitioning territory. We proposed a technique based on partitioning territory with the use of hand shape gestures.

7.3 Window Avatar Technique

In this section, we described Window Avatar, which allow the user to create a window to enhance interactions in multi-touch tabletops based on WIMP paradigm. We also proposed a set of interaction techniques based on Window Avatar, and classified them into two categories: manipulation of Window Avatar and collaboration through Window Avatar (see Table 7. 1).

Manipulation of Window Avatar	Cooperation through Window Avatar
Creating Window Avatar	Window Avatar Connection
Add and Remove Objects	Object Transfer and Window Sharing
Move, Rotate and Zoom Window	Access Remote Territory
Multi Layered Windows	
Group and Store Objects	

Table 7. 1 Techniques of window avatar.

7.3.1 Manipulation of Window Avatar

Creating Window Avatar

The creation of Window Avatar depends on detecting the vocabulary of two hand shapes: Tile and Curved Rail. Tile is a hand shape that the user makes by placing the palm on the table, with the thumb spreading apart. Curved Rail is a curved hand pose with a curved angle of around 90 degrees. In our prototype, these hand shapes were recognized simply by examining the eccentricity and the size of the ellipse: a rounded shape detected as Tile, and curved thin shape as Rail. While simple, his eccentricity-based detection works reliably in our prototype; however, more elaborate solutions might be necessary if greater robustness is desired.

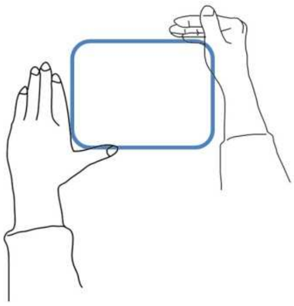


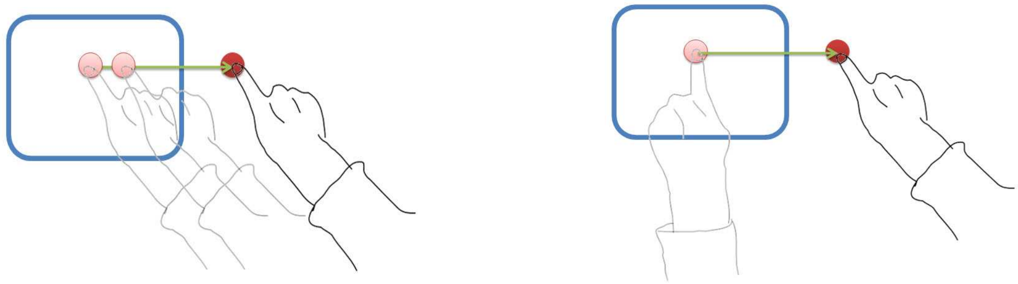
Figure 7. 1 Window avatar.

If the system detects Tile and Curved Rail, a Window Avatar is created (see Figure 7. 1 Window avatar.). Window Avatar is a rectangular area which is determined by the contact position of Tile and Curved Rail; a vertex of the rectangle is the center point within the area of Tile, and the opposite vertex is the center point within the area of Curved Rail. The user can create more than one Window Avatar.

The area within the window is a privacy territory. The user can manipulate objects in this area, while other users can not access objects within this area without permission.

Adding and Removing Objects

The user can employ two ways to remove objects from a window. First, as illustrated in Figure 7. 2a, the user can click on an object with a finger and then drag it out of the window. Second, as illustrated in Figure 7. 2b, the user can click on an object with a finger and click a position outside the window with another finger. A new object will appear in the second clicked position, along with a line linking these two objects, indicating the new object is a copy of the old object. Any manipulation on the new object will lead to the same manipulation on the old object. The user can perform a slide gesture on the line to disconnect the two targets. Similarly, the user can add objects from a window by dragging them into the window or by clicking on an object and then clicking a position inside the window.

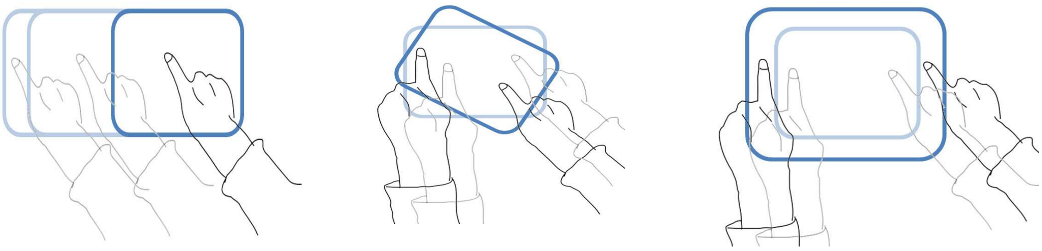


(a) (b)

Figure 7. 2 Remove objects.

Moving, Rotating and Zooming Window

The user can translate, rotate and zoom a window to adjust its position and orientation. By placing one or more fingers on the background of the window (i.e. anywhere objects are not located) and dragging it, the user can move the window (see Figure 7. 3a). The window can be rotated by placing one finger on the background and rotating another finger around the finger on the screen (see Figure 7. 3b). Pinching together and spreading apart two fingers are used to zoom out and zoom in the window respectively (see Figure 7. 3c). In summary, Window Avatar is moveable, scalable and reoriented, allowing the user to adjust the configurations of the window for better interactions.



(a) (b) (c)

Figure 7. 3 (a) Move, (b) rotate and (c) zoom window.

Multi Layered Windows

Window Avatar can provide various layers to present different groups of information. As shown in Figure 7. 4, the user can perform flick gesture on the bars at the top and bottom to switch different layers. Two layers are initiated when the Window Avatar is created; one presents the territory covered by Window Avatar, the other presents the entire screen area. The user can also add a layer by performing flick gesture on the bars at the left and right sides.

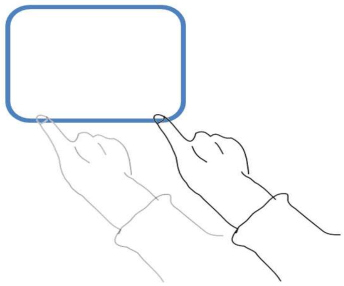


Figure 7. 4 Window switch.

Grouping and Storing Objects

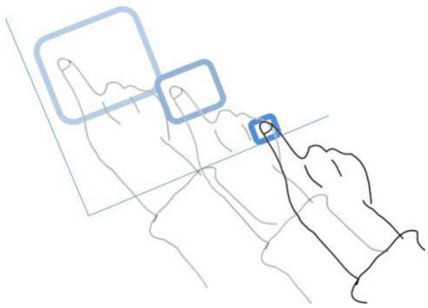
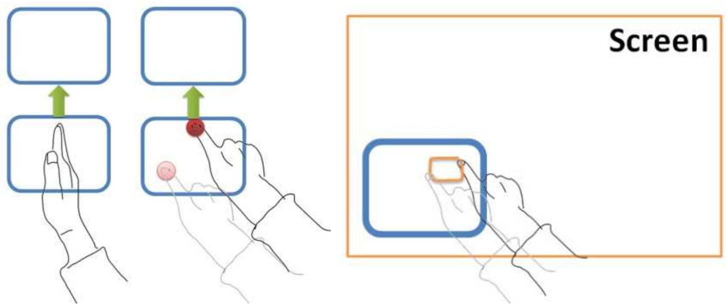


Figure 7. 5 Store objects.

In the daily life, people usually use folders to sort documents, so as to better retrieve information. Users often perform a variety of interactive activities such as object classification in multi-touch tabletops, hence it is important to design a technique to group and store objects. Our technique, Window Avatar, can enable the user to create multi windows and sort objects into different windows. The user can also store a window by moving the window to the bottom of the surface (see Figure 7. 5). An icon representing the window will be shown at the bottom of the surface. To view the window in original size, the user needs to place a finger on the icon and move it towards the center of the screen.

7.3.2 Collaboration through Window Avatar

An important advantage of Window Avatar is that it can support collaborative interactions. In this section, we will introduce how users collaborate by means of Window avatar.



(a) (b) (c)

Figure 7. 6 (a) Window avatar connection. (b) Object transfer and window sharing. (c) Access remote territory on the screen.

Window Avatar Connection

A window Avatar can be connected and disconnected to other Window Avatars. As illustrated in Figure Figure 7. 6a, to connect another Window Avatar, the user should perform a Rail gesture, which is detected as a thin long shape. Then a line with an arrow will be shown, representing the link between two windows. The orientation of the hand shape indicates which remote Window Avatar will be connected to local Window Avatar. This hand gesture is easy for users to remember and use, because in people's daily life, the use of this hand gesture can help people to determine orientation. To disconnect another Window Avatar, the user needs to perform a slide gesture on the line.

Object Transfer and Window Sharing

Window Avatar allows the exchange of digital objects among multiple users. After setting up a connection with a window avatar created by another user, the user can easily send an object by dragging it into the connection line (see Figure 7. 6b). Then the object will be shown in the target window. In addition, the user can share the whole window of Window Avatar with other users by placing one finger in the background of the Window Avatar and dragging it into the connection line.

Access Remote Territory on the Screen

Window Avatar provides a layer in which the user can view the entire screen. In this layer, the user can select the interested territory of the screen by drawing a closed curve in the territory (see Figure 7. 6c), or by double tapping the center area of the territory. Then the territory will be enlarged and be presented in the area of the Window Avatar. In this way, the user can view the remote territory and manipulate remote objects in the territory.

7.4 Discussion

By means of Window Avatar, users can perform individual and collaborative activities in multi-touch tabletop displays. Regarding individual activities, the user can aggregate and manipulate groups of objects based on Window Avatar. In addition, the user can translate, rotate and scale a group of digital targets in Window Avatar with a single gesture consisting of simultaneous movement of two fingers. The ability to store resource items anywhere in the workspace and move them around can be important for collaborative task and group interactions on a table. Existing storage techniques, such as Storage Bins [86] and TableTrays [74] were designed for the pen-based digital workspace, hence may not suit for multi-touch tabletop displays. Window Avatar, which is designed based on the properties of multi-touch, provides a suitable tool for organizing and sharing information for multi-touch based interactions.

On the other hand, for user collaborative activities on multi-touch tabletop displays, the large size and horizontal orientation of the display lead to certain challenges for designing effective collaborative user interfaces. Window Avatar can support information access and sharing for multi users on a tabletop workspace, so as to enhance users' sense of teamwork, increase awareness of important system events, facilitate reachability and access control on large and shared displays.

7.5 Conclusion

We proposed Window Avatar, a new interaction technique to better support individual and collaborative activities in multi-touch tabletop displays. This technique incorporates many of the capabilities that have been proposed for tabletop groupware into a single mechanism. It allows users to temporarily group sets of objects and organize the table area, and also to support information access and sharing for multi users on a tabletop workspace. As multi-touch tabletops are gaining popularity in collaborative activities, it is important to design interaction techniques which can improve interactions on manipulation and collaboration. Our work is one step in this exploration.

Chapter 8 General Conclusions and Future Research Directions

8.1 Summary of Contributions

Gesture-based interaction has been widely employed in the design of touch screen devices. This interaction style can provide users natural and convenient operation feeling, hence drawing much attention in HCI research field. This thesis presents five studies which focused on two important issues regarding touch-based gesture interactions. Three of them are referred to how to design touch-based gestures, with regard to different input forms (pen vs. finger), users of different ages (older users vs. younger users), and different entry sizes. The other two are about how to employ touch-based gestures in interactive activities: gesture performance in a document scrolling task of touch-based mobile phones and the use of touch gestures to better support multi-user collaborative tasks on large tabletops respectively.

For gesture-oriented design, three studies were conducted to improve touch based gesture design. The purpose of the first study is to quantify the differences and similarities between finger and pen gestures. The work has provided a methodology to investigate and quantify the performance of finger and pen gestures, in which finger and pen gestures were analyzed according to multiple features that characterize stroke gestures. Some features revealed similarities between finger and pen drawn gestures; finger gesture design should exploit these features but avoid using the features which were less accurate with the finger than with the pen. This work provides a solid foundation to apply principles, methods and findings from pen-based gesture design to finger-based gesture design. Second, a user-defined gesture study was conducted to compare user-defined gestures between younger people and older people in the context of pen input and finger input. It was found that (1) gesture design should avoid using gestures with high Degree of Freedom for older people; (2) desktop paradigm has less effect on gesture performance for old people than for younger people. (3) analogue gestures are easy to remember and use for older people, hence should be deeply explored and widely used. The study demonstrates that understanding the preferred gestures of both younger adults and older adults is important for gesture design. Third, as gesture entry size is an important factor for determining users' performance of gesture input, a study was presented to quantitatively investigate optimal finger-based entry size in touch-based mobile phones for two commonly used Chinese handwriting input styles: two-handed entry with the non-dominant hand holding the device and the index finger of the dominant hand entering characters; and one-handed entry with the dominant hand holding the device and the thumb of the dominant hand being used for character entry. Results were assessed in terms of the number and length of protruding strokes, writing time, stroke writing speed, size ratio, number of writing attempts and subjective preference. For both one-handed entry and two-handed entry, the optimal entry box size for handwritten Chinese characters was found to be 2.5cm×2.5cm. This size entry box is large enough for fast and accurate handwriting with high entry area utilization rate and few, short protruding strokes. The experimental results and methodology of this study can be employed in user interface design for gesture-based interaction in touch-based mobile phones.

With respect to gesture-based task, two studies were conducted to analyze gesture-based task and support gesture-based interaction. First, because flick and ring are two important scrolling techniques for document navigation, examining the advantages and disadvantages of these two scrolling techniques would be beneficial to scrolling technique design. For this purpose, this thesis quantitatively analyzed the performance of two scrolling techniques (flick and ring) for document navigation in touch-based mobile phones by means of three input methods (index finger, pen and thumb), with specific consideration given to two motor activities: sitting and walking. Our findings were as follows: (1) overall, in both sitting and walking activities, for the three input methods, flick resulted in shorter movement time and fewer numbers of crossings than ring, suggesting that flick is superior to ring for document navigation in touch-based mobile phones; (2) for sitting activity, regarding pen and thumb input, there were interaction effects between scrolling technique and target distance. Ring led to shorter movement time than flick for large target distance. This finding indicates that ring has a potential interaction advantage, which should be deeply explored for future scrolling technique design; (3) regarding sitting and walking activities, both flick and ring document scrolling in touch-based mobile phones can be modeled by the Anderson's model [2]. We believe these findings offer several insights for scrolling technique design for document navigation in touch-based mobile phones. Second, this thesis proposed Window Avatar, a window-based technique which allows the user to create a personal territory by means of hand shape gestures in multi-touch tabletop displays. Based on Window Avatar, a set of interaction techniques were presented using shape gestures in combination with direct manipulations, so as to enhance user interaction on manipulation and collaboration. The user can aggregate and manipulate groups of objects with these interaction techniques. In addition, these interactions can provide users to better coordinate access to space and resources, aid users in view remote territories on large displays, and result in a high level of group awareness as well.

In summary, this dissertation contributes to the field of gesture-based interaction in view of gesture-oriented design and gesture-based tasks. The research topics studied here plays an important role in touch-based gesture design. As a basic understanding of these topics has been established, further research issues are exposed, which should be pursued in future studies.

8.2 Future Research Directions

As introduced in the section of Introduction, gesture related research covers wide range of research topics in HCI field, such as gesture-related models, gesture recognition, and feedback of gesture input. This thesis focused on two topics of touch-based gesture interaction: (1) gesture-oriented design, with specific consideration of input forms, users ages and entry sizes; (2) gesture-based task, which aims at a document scrolling task in touch-based mobile phones and a collaborative task on large tabletops. These studies provide a number of valuable findings and methodologies for future gesture research regarding other aspects of gesture-based interaction.

This thesis examined the differences and similarities of pen gesture and finger gesture in a Tablet PC in the context of sitting posture for younger adults. Future work will concern the investigation of pen and finger gesture in mobile devices in the context of walking posture.

This thesis investigated the gesture performance in a document scrolling task in touch-based mobile phones and a collaborative task on large tabletops. In the future, we seek to employ gesture-based interactions in other interaction tasks such as target selection and examined the gesture performance in comparison with other input techniques.

Acknowledgement

Foremost, I would like to express my deep and sincere gratitude to my supervisor Prof. Xiangshi Ren for the great guidance of my Ph.D study and research. His high motivation and strong enthusiasm to HCI research, as well as his immense knowledge of HCI field, have a deep impact on me, and undoubtedly will have a far-reaching influence on my future research and work. His understanding, encouraging and personal guidance have provided a good basis for the present thesis.

My sincere thanks also go to Dr. Shumin Zhai. His thoughts and guidance promoted an original idea to the work published in CHI 2012. I am also very grateful to Prof. Inoue, Prof. Hamamura, Prof. Hoshino and Prof. Shimizu for the enlightening discussion and insightful comments on my thesis. Especially, I appreciate Xindong Yang' help when I started my study in HCI field.

I owe my loving thanks to my wife, Mrs. Yanan Zhang, who offered me great supports in the past three years of Ph.D study, not only regarding delicious food cooked for me but also about spiritual encouragement and support.

Last but not the least, I would like to thank the members of the Ren lab, particularly Yizhong Xin, Feng Wang, Minghui Sun, Yusuke Hayashi, Masatoshi Kusuba, Yuan Fu, Yang Gao, for their kind help with my work.

During my Ph.D study, I was supported by the scholarships provided by Kochi University of Technology, JASSO scholarship and travel grant provided by NEC C&C Research. The financial support is gratefully acknowledged.

References

[1]Aliakseyeu, D., Irani, P., Lucero, A. and Subramanian, S. Multi-flick: an evaluation of flick-based scrolling techniques for pen interfaces. In *Proc. CHI 2008*, ACM Press (2008), 1689-1698.

[2]Andersen, T.H. A simple movement time model for scrolling. *Ext. abstracts. CHI 2005*, ACM Press (2005), 1180-1183.

[3]Anderson, D., Bailey, C. and Skubic, M. Hidden Markov Model symbol recognition for sketch-based interfaces. *AAAI Fall Symposium 2004*, AAAI Press (2004), 15-21.

[4]Andersen, T.H. and Zhai, S. "Writing with music": exploring the use of auditory feedback in gesture interfaces. ACM *TAP* 7, 3(2010), 1-24.

[5]Appert, C. and Zhai, S. Using strokes as command shortcuts: cognitive benefits and toolkit support. In *Proc. CHI 2009*, ACM Press (2009), 2289-2298.

[6]Bau, O. and Mackay, W. E. OctoPocus: a dynamic guide for learning gesture-based command sets. In *Proc. UIST 2008*, ACM Press (2008), 37-46.

[7]Bauer, D., Fastrez, P. and Hollan, J. Computationally-enriched "piles" for managing digital photo collections. In *Proc. VLHCC 2004*, IEEE Computer Society Press (2004), 193-195.

[8]Blagojevic, R., Chang, S.H. and Plimmer, B. The power of automatic feature selection: Rubine on steroids. In *Proc. SBIM 2010*, Eurogaphics Association (2010), 79-86.

[9]Bragdon, A., Zeleznik, R., Williamson, B., Miller and T., Jr. J. GestureBar: improving the approachability of gesture-based interfaces. In *Proc. CHI 2009*, ACM Press (2009), 2269-2278.

[10]Bragdon, A. and Ko, H.S. Gesture select: acquiring remote targets on large displays without pointing. In *Proc. CHI 2011*, ACM Press (2011), 187-196.

[11]Bragdon, A., Uguray, A., Wigdor, D., Anagnostopoulos, S., Zeleznik, R. and Feman, R. Gesture play: motivating online gesture learning with fun, positive reinforcement and physical metaphors. In *Proc. CHI 2010*, ACM Press (2010), 39-48.

[12]Cao, X. and Balakrishnan, R. Evaluation of an on-line adaptive gesture interface with command prediction. In *Proc. GI 2005*, ACM Press (2011), 187-194.

[13]Cao, X., Wilson, A.D., Balakrishnan, R., Hinckley, K. and Hudson, S.E. ShapeTouch: leveraging contact shape on interactive surfaces. In *Proc. ITS 2008*, ACM Press (2008), 129-136.

[14]Cao, X. and Zhai, S. Modeling human performance of pen stroke gestures. In *Proc. CHI 2007*, ACM Press (2007), 1495-1504.

[15]Chatty, S. and Lecoanet, P. Pen computing for air traffic control. In *CHI 1996,* ACM Press (1996), 87-94.

[16]Chiu, P., Liao, C. and Chen, F. Multi-touch document folding: gesture models, fold directions and symmetries. In *CHI 2011,* ACM Press (2011), 1591-1600.

[17]Cho, M.G. A new gesture recognition algorithm and segmentation method of Korean scripts for gesture-allowed ink editor. *Information Sciences*, *176*, 9(2006), 1290-1303.

[18]Cockburn, A., Gutwin, C., Alexander, J. Faster Document Navigation with Space-Filling Thumbnails. In *Proc. CHI 2006,* ACM Press (2006), 1-10.

[19]Cockburn, A., Savage, J., and Wallace, A. Tuning and testing scrolling interfaces that automatically zoom. In *Proc. CHI 2005,* ACM Press (2005), 71-80.

[20]Commarford, P.M. An investigation of text throughput speeds associated with Pocket PC input method editors. *International Journal of Human-Computer Interaction*, 17, 3(2004), 293-308.

[21]Diehl, J., Mollers, M. and Borchers, J. Improving list selection performance with pressure-sensitivity on a scroll ring. *Ext. abstracts. UIST 2008*, ACM Press (2008), 65-72.

[22]DuoSense Digitizer Product Brief. http://www.n-trig.com.

[23]Findlater, L., Lee, B. and WObbrock, J. Beyond QWERTY: augmenting touch screen keyboards with multi-touch gestures for non-alphanumeric input. In *Proc. CHI 2012*, ACM Press (2012), 2679-2682.

[24]Freeman, D., Benko, H., Morris, M.R. and Wigdor, D. ShadowGuides: visualizations for in-situ learning of multi-touch and whole-hand gestures. In *Proc. ITS 2009*, ACM Press (2009), 165-172.

[25]Goldberg, D. and Richardson, C. Touching-typing with a stylus. In *Proc. CHI 1993*, ACM Press (1993), 80-87.

[26]Hinckley, K., Baudisch, P., Ramos, G. and Guimbretiere, F. Design and analysis of delimiters

for selection-action pen gesture phrases in scriboli. In *Proc. CHI 2005*, ACM Press (2005), 451-460.

[27]Hinckley, K., Cutrell, E., Bathiche, S. and Muss, T. Quantitative analysis of scrolling techniques. In *Proc. CHI 2002*, ACM Press (2002), 65-72.

[28]Hinckley, K., Yatani, K., Pahud, M., Coddington, N., Rodenhouse, J., Wilson, A., Benko, H. and Buxton, B. Pen + Touch = New Tools. In *Proc. UIST2010*, ACM Press (2010), 27-36.

[29]Hirotaka, N. Reassessing current cell phone designs: using thumb input effectively. *Ext. abstracts. CHI 2003*, ACM Press (2003), 938-939.

[30]Hong, J. I. and Landay, J. A. SATIN: a toolkit for informal ink-based applications. In *Proc. UIST2000*, ACM Press (2000), 63-72.

[31]Hourcade, J.P. and Berkel, T.R. Simple pen interaction performance of young and older adults using handheld computers. *Interacting with Computers*, 20(2008), 166-183.

[32]Hourcade, J.P., Nguyen, C.M., Perry, K.B. and Denburg, N.L. PointAssist for older adults: analyzing sub-movement characteristics to aid in pointing tasks. In *Proc. CHI 2010*, ACM Press (2010), 1115-1124.

[33]HP TouchSmart tx2 Notebook PC Maintenance and Service Guide. http://www.hp.com.

[34]Hwang, F., Batson, H. and Williams, N. Bringing the target to the cursor: Proxy targets for older adults. *Ext. abstracts. CHI 2008*, ACM Press (2008), 2775-2780.

[35]Igarashi, T. and Hinckley, K. Speed-dependent automatic zooming for browsing large documents. In *Proc. UIST 2000*, ACM Press (2000), 139-148.

[36]Isokoski, P. Model for unistroke writing time. In *Proc. CHI 2001*, ACM Press (2001), 357-364.

[37]Kammer, D., Wojdziak, J., Keck, M., Groh, R. and Taranko, S. Towards a formalization of multi-touch gestures. In *Proc. ITS 2010*, ACM Press (2010), 49-58.

[38]Karlson, A.K., Bederson, B.B. and Contreras-Vidal, J.L. Understanding single-handed mobile device interaction. *Tech Report HCIL-2006-02*.

[39]Khandkar, S.H., Sohan, S.M. and Sillito, J. Tool support for testing complex multi-touch gestures. In *Proc. ITS 2010*, ACM Press (2010), 59-68.

[40]Kin, K., Hartmann, B. and Derose, T. Proton: multitouch gestures as regular expressions. In *Proc.*

CHI 2012, ACM Press (2012), 2885-2894.

[41]Klinkhammer, D., Nitsche, M., Specht, M. and Reiterer, H. Adaptive personal territories for co-located tabletop interaction in a museum setting. In *Proc. ITS 2011*, ACM Press (2011), 107-110.

[42]Knoblauch, R. L., Pietrucha, M.T. and Nitzburg, M. Field studies of pedestrian walking speed and start-up time. *Transportation Research Record 1538*, 1996, 27-38.

[43]Kray, C. Nesbitt, D., Dawson, J. and Rohs, M. User-defined gestures for connecting mobile phones, public displays, and tabletops. In *Proc. MobileHCI 2010*, ACM Press (2010), 239-248.

[44]Kristensson, P.O. and Denby, L.C. Text entry performance of state of the art unconstrained handwriting recognition: A longitudinal user study. In *Proc.CHI 2009*, ACM Press (2009), 567-570.

[45]Kristensson, P.O. and Zhai, S. Command strokes with and without preview: using pen gestures on keyboard for command selection. In *Proc.CHI 2007*, ACM Press (2007), 1137-1146.

[46]Kristensson, P.O. and Zhai, S. SHARK2: a large vocabulary shorthand writing system for pen-based computers. In *Proc. UIST 2004*, ACM Press (2004), 43-52.

[47]Kruger, R., Carpendale, S., Scott, S.D. and Tang, A. Fluid integration of rotation and translation. In *Proc. CHI 2005*, ACM Press (2005), 601-610.

[48]Kurihara, K., Goto, M., Ogata, J. and Igarashi, T. Speech pen: predictive handwriting based on ambient multimodal recognition. In *Proc. CHI2006*, ACM Press (2006), 851-860.

[49]Kurtenbach, G. and Buxton, W. Issues in combining marking and direct manipulation techniques. In *Proc. UIST 1991*, ACM Press (1991), 137-144.

[50]Kurtenbach, G., Sellen, A. and Buxton, W. An empirical evaluation of some articulatory and cognitive aspects of “marking menus”. *Human Computer Interaction*, 8, 1 (1993), 1-23.

[51]Kurtenbach, G. and Hulteen, E. Gestures in Human-computer communications. In B. Laurel (Ed.) *The Art of Human Computer Interface Design*. Addison-Wesley (1990), 309-317.

[52]Lacuaniti, F., Terzuolo, C. and Viviani, P. The law relating the kinematic and figural aspects of drawing movements. *Acta Psychologica*, 54 (1983), 115-558.

[53]Lahey, B. Girouard, A., Burleson, W. and Vertegaal, R. PaperPhone: Understanding the Use of

Bend Gestures in Mobile Devices with Flexible Electronic Paper Displays. In *Proc. CHI 2011*, ACM Press (2011), 1303-1312.

[54]Landay, J. A. and Myers, B. A. Interactive sketching for the early stages of user interface design. In *Proc. CHI 1995*, ACM Press (1995), 43-50.

[55]Lee, S., Kim, S., Jin, B., Choi, E., Kim, B., Jia, X., Kim, D. and Lee, K. How users manipulate deformable displays as input devices. In *Proc. CHI 2010,* ACM Press (2010), 1647-1656.

[56]Li, Y. Gesture search: a tool for fast mobile data access. In *Proc.* UIST 2010, ACM Press (2010), 87-96.

[57]Li, Y. Protractor: A fast and accurate gesture recognizer. In *Proc.* CHI 2010, ACM Press (2010), 2169-2172.

[58]Liao, C., Guimbretiere, Hinckley, K. and Hollan, J. Papiercraft: A gesture-based command system for interactive paper. *ACM Transactions on Computer Human Interaction 14*, 4 (2008), 1-27.

[59]Long, A. C., Landay, J. A. and Rowe, L. A. Implications for a gesture design tool. In *Proc. CHI 1999*, ACM Press (1999), 40-47.

[60]Long, A.C., Landay, J.A., Rowe, L.A. and Michiels, J. Visual similarity of pen gestures. In *Proc. CHI 2000,* ACM Press (2000), 360-367.

[61]Lü, H. and Li, Y. Gesture Avatar: a technique for operating mobile user interfaces using gestures. In *Proc. CHI 2011*, ACM Press (2011), 207-216.

[62]MacKenzie, I.S., Nonnecke, B., Riddersma, S., McQueen, C. and Meltz, M. Alphanumeric entry on pen-based computers. *International Journal of Human-Computer Studies*, 41, 1994, 775–792.

[63]Moffatt, K. and McGrenere, J. Steadied-bubbles: combining techniques to address pen-based pointing errors for younger and older adults. In *Proc. CHI 2010*, ACM Press (2010), 1125-1134.

[64]Morris, M.R., Huang, A., Paepcke, A. and Winoqrad, T. Cooperative gestures: multi-user gestural interactions for co-located groupware. In *Proc. CHI 2006*, ACM Press (2006), 1201-1210.

[65]Moscovich, T. and Hughes, J.F. Navigating documents with the virtual scroll ring. In *Proc. UIST 2004*, ACM Press (2004), 65-72.

[66]Myers, C.S. and Rabiner, L.R. A comparative study of several dynamic time-warping algorithms for connected word recognition. *The Bell System Technical J., 60*, 7(1981), 1389-1409.

[67]Newman, M. W., Lin, J., Hong, J. I. and Landay, J. A. DENIM: an informal web site design tool inspired by observations of practice. *Human Computer Interaction, 18*, 3(2003), 259-324.

[68]Newman, W. M. and Sproull, R. F. Principles of Interactive Computer Graphics. New York: McGraw-Hill Press (1979).

[69]Notowidigdo, M. and Miller, R.C. Off-line sketch interpretation. *AAAI Fall Symposium. Menlo Park*, CA: AAAI Press (2004), 120-126.

[70]Natural User Interface. http://en.wikipedia.org/wiki/Natural_user_interface

[71]Ouyang, T.Y. and Li, Y. Bootstrapping Personal Gesture Shortcuts with the Wisdom of the Crowd and Handwriting Recognition. In *Proc. CHI 2012*, 2895-2904.

[72]Parhi, P., Karlson, A.K and Bederson, B.B. Target size study for one-handed thumb use on small touchscreen devices. In *Proc. MobileHCI 2006*, ACM Press (2006), 203–210.

[73]Pedersen, E. R., McCall, K., Moran, T. P. and Halasz, F.G. Tivoli: an electronic whiteboard for informal workgroup meetings. In *Proc. INTERACT 1993*, 391-398.

[74]Pinelle, D., Stach, T., Gutwin, C. TableTrays: temporary, reconfigurable work surfaces for tabletop groupware. In *Proc. TABLETOP 2008*, IEEE Computer Society Press (2008), 41-48.

[75]Pittman, J.A. Recognizing handwritten text. In *Proc. CHI 1991*, ACM Press (1991), 271-275.

[76]Rekimoto, J. SmartSkin: an infrastructure for free-hand manipulation on interactive surfaces. In *Proc. CHI 2002*, ACM Press (2002), 113-120.

[77]Ren, X. and Moriya, S. Improving selection performance on pen-based systems: a study of pen-based interaction for selection tasks. *TOCHI, 7*, 3(2000), 384-416.

[78]Ren, X., Tamura, K., Kong, J. and Zhai, S. Candidate display styles in Japanese input. In *Proc. INTERACT 2003*, 868-871.

[79]Ren, X. and Zhou, X. The optimal size of handwriting character input boxes on PDAs, *Int. J. Human-Computer Interaction*, 25, 8 (2009), 762-784.

[80]Rubine, D. Specifying gestures by example. In *Proc. SIGGRAPH 1991*, ACM Press (1991), 329-337.

[81]Ruiz, J., Li, Y. and Lank, E. User-defined motion gestures for mobile interaction. In *Proc. CHI 2011*, ACM Press (2011), 403-412.

[82]Sae-Bae, K., Ahmedm, K., Isbister, K. and Memon, N. Biometric-rich gestures: a novel approach to authentication on multi-touch devices. In *Proc. CHI 2012*, ACM Press (2012), 403-412.

[83]Sarcher, H., Tng, T.H. and Loudon, G. Beyond translation: Approaches to interactive products for Chinese consumers. *International Journal of Human Computer Interaction*, 13, 1(2001), 41–51.

[84]Schraefel, M., Smith, G. M. and Baudisch, P. Curve dial: eyes-free parameter entry for GUIs. *Ext. abstracts. CHI 2005*, ACM Press (2005), 1146-1147.

[85]Schmidt, R. A. Motor control and learning: a behavioral emphasis. 2nd ed. *Champaign*, IL: Human Kinetics (1988).

[86]Scott, S.D., Carpendale, M.S., Habelski, S. Storage bins: mobile storage for collaborative tabletop displays. In *Proc. CGA 2005*, IEEE Computer Society Press (2005), 58-65.

[87]Scott, S.D., Sheelagh, M., Carpendale, T. and Inkpen, K.M. Territoriality in collaborative tabletop workspaces. In *Proc. CSCW 2004*, ACM Press (2004), 294-303.

[88]Sears, A., Revis, D., Swatski, J., Crittenden, R. and Shneiderman, B. Investigating touchscreen typing: The effect of keyboard size on typing speed. *Behaviour and Information Technology*, 12, 1993, 17–22.

[89]Sears, A. and Zha, Y. Data entry for mobile devices using soft keyboards: Understanding the effects of keyboard size and user tasks. *International Journal of Human-Computer Interaction*, 16, 2(2003), 163-184.

[90]Sezgin, T.M. and Davis, R. HMM-based efficient sketch recognition. In *Proc. IUI* 2005, ACM Press (2005), 281-283.

[91]Shen, C., Vernier, F., Forlines, C. and Ringel, M. DiamondSpin: an extensible toolkit for around-the-table interaction. In *Proc. CHI 2004*, ACM Press (2004), 167-174.

[92]Shilman, M., Tan, D. S. and Simard, P. CueTIP: A mixed-initiative interface for correcting handwriting errors. In *Proc. UIST 2006*, ACM Press (2006), 323–332.

[93]Smith, G.M. and Schraefel, M.C. The radial scroll tool: scrolling support for stylus- or

touch-based document navigation. In *Proc. UIST 2004*, ACM Press (2004), 53-56.

[94]Smith, K. U. and Smith, T. J. A study of handwriting and its implications for cognitive considerations in Human-Computer Interactions. *International Journal of Human-Computer Interaction*, 3, 1(1991), 1-30.

[95]Sun, L., Guimbretière, F. Flipper: a new method of digital document navigation, *Ext. abstracts. CHI 2005*, ACM Press (2005), 2001-2004.

[96]Sutherland, I. E. Sketchpad: A man-machine graphical communication system: MIT Lincoln Laboratory, *Technical Report (1963)*, No. 296.

[97]Tappert, C.C. Cursive script recognition by elastic matching. *IBM J. of Research & Development*, *26*, 6(1982), 765-771.

[98]Tashman, C. and Edwards, K. LiquidText: a flexible, multitouch environment to support active reading. In *Proc. CHI 2011,* ACM Press (2011), 3285-3294.

[99]Tse, E., Histon, J., Scott, S.D and Greenberg, S. Avoiding interference: how people use spatial separation and partitioning in SDG workspaces. In *Proc. CSCW 2004*, ACM Press (2004), 252-261.

[100]Tse, E., Greenberg, S., Shen, S., Barnwell, J., Shipman, S. and Leigh, D. Multimodal split view tabletop interaction over existing applications. In *Proc. TABLETOP 2007*, IEEE Computer Society Press (2007), 129-136.

[101]Viviani, P. and Flash, T. Minimum-jerk, two-thirds power law, and isochrony: converging approaches to movement planning. *Journal of Experimental Psychology: Human Perception and Performance, 21*, 1 (1995), 32-53.

[102]Viviani, P. and Terzuolo, C. Trajectory determines movement dynamics. *Neuroscience, 7*, 2 (1982), 431-437.

[103]Wigdor, D., Benko, H. and Pella, J. Rock & rails: extending multi-touch interactions with shape gestures to enable precise spatial manipulations. In *Proc. CHI 2011*, ACM Press (2011), 1581-1590.

[104]Wright, C. E. Generalized motor programs: reexamining claims of effector independence in writing. In *Attention and Performance XIII, Erlbaum*, Hillsdale, NJ (1990), 294-320.

[105]Wobbrock, J., Morris, M. and Andrew, A. User-defined gestures for surface computing. In *Proc. CHI 2009,* ACM Press (2009), 1083-1092.

[106]Wobbrock, J. O., Myers, B. A. and Chau, D. H. In-stroke word completion. In *Proc. UIST 2006*, ACM Press (2006), 333–336.

[107]Wobbrock, J. O., Myers, B. A. and Kembel, J. A. EdgeWrite: A stylus-based text entry method designed for high accuracy and stability of motion. In *Proc. UIST 2003,* ACM Press (2003), 61–70.

[108]Wobbrock, J. O., Wilson, A. D. and Li, Y. Gestures without libraries, toolkits or training: a $1 recognizer for user interface prototypes. In *Proc. UIST 2007,* ACM Press (2007), 159-168.

[109]Wherry, E. Scroll ring performance evaluation. *Ext. abstracts. CHI 2003*, ACM Press (2003), 758-759.

[110]Wilson, A.D. and Shafer, S. XWand: UI for intelligent spaces. I*n Proc. CHI 2003*. ACM Press (2003), 545-552.

[111]Wolf, C. G. and Rhyne, J. R. Gesturing with shared drawing tools. I*n Proc. CHI 1993*. ACM Press (1993), 137-138.

[112]Wu, M. and Balakrishnan, R. Multi-finger and whole hand gestural interaction techniques for multi-user tabletop displays. In *Proc. UIST 2003*, ACM Press (2003), 193-202.

[113]Wu, M., Shen, C., Ryall, K., Forlines, C. and Balakrishnan, R. Gesture Registration, Relaxation, and Reuse for Multi-Point Direct-Touch Surfaces. *In Proc. First IEEE International Workshop on Horizontal Interactive Human-Computer Systems 2006*, IEEE Computer Society Press (2006), 185-192.

[114]Yin, J. and Ren, X. ZWPS and pressure scroll: two pressure-based techniques in pen-based interfaces. *IPSJ*, 3(2007), 767-778.

[115]Zhai, S. and Kristensson, P.O. Shorthand writing on stylus keyboard. In *Proc. CHI 2003*, ACM Press (2003), 97-104.

[116]Zhai, S., Kristensson, P.O, Gong, P., Greiner, M., Peng, S.A., Liu, L.M. and Dunnigan, A. Shapewriter on the iphone: from the laboratory to the real world. *Ext. abstracts. CHI 2009*, ACM Press (2009), 2667-2670.

[117]Zhai, S., Smith, B. A., and Selker, T. Improving browsing performance: A study of four input devices for scrolling and pointing tasks. In *Proc. INTERACT 1997*, Chapman & Hall, Ltd. Press (1997), 286-293.

[118]Zhou, X., Zhao, S., Chignell, M. and Ren, X. Assessing age-related performance decrements in user interface tasks. In *Proc. ICIA 2011*, IEEE Press (2011), 817-822.

Appendix A Publications

A.1 Articles in Refereed Journals

1. Tu, H., Wang, F., Tian, F. and Ren, X. (2012). A Comparison of Flick and Ring Document Scrolling in Touch-based Mobile Phones, accepted in *International Journal of Human-Computer Interaction*.

2. Tu, H. and Ren, X. (2012). Optimal Entry Size of Handwritten Chinese Characters in Touch-based Mobile Phones, to appear in *International Journal of Human-Computer Interaction*.

3. Tu, H. and Ren, X. (2011). Finger Chording in the Air, *Innovative Computing, Information and Control Express Letters, 6*, 6(2011), 1623-1628.

A.2 Articles in Full Paper Refereed International Conference Proceedings

1. Tu, H., Ren, X. and Zhai, S. (2012). A Comparative Evaluation of Finger and Pen Stroke Gestures, in *Proceedings of CHI 2012*, ACM Press (2012), 1287-1296, Acceptance rate of 23%.

2. Tu, H., Yang, X., Wang, F., Tian, F. and Ren, X. (2012). Experimental Evaluation of Different Mode Switching Techniques in Pen-based Handheld Devices, accepted in *Proceedings of APCHI 2012*. Acceptance rate of 26.5%.

3. Tu, H., Wang, F., Tian, F. and Ren, X. (2012). A Comparison of Flick and Ring Document Scrolling in Touch-based Mobile Phones, accepted in Proceedings of APCHI 2012 and recommended to *International Journal of Human-Computer Interaction*. Acceptance rate of 26.5%.

4. Tu, H. and Ren, X. (2011). Finger Chording in the Air, accepted in *Proceedings of ICICIC 2011*.

A.3 Articles in Refereed International Conference Proceedings

1. Mizobata, R., Tu, H. and Ren, X. (2012). User-defined Motion Gestures, accepted in *Proceedings of APCHI 2012*.

2. Hayashi, Y., Tu, H. and Ren, X. (2012). An Empirical Investigation into Differences and Similarities between Age-related Stroke Gestures, accepted in *Proceedings of APCHI 2012*.

3. Okamoto, M., Tu, H. and Ren, X. (2012). A Comparative Evaluation of Finger and Pen Stroke Gestures in Mobile Environments, accepted in *Proceedings of APCHI 2012*.

4. Kusuba, M., Tu, H. and Ren, X. (2012). User-defined Gestures for Older People, accepted in *Proceedings of APCHI 2012*.

A.4 Articles in Non-refereed Local Conference Proceedings

1. Mizobata, R., Tu, H. and Ren, X. (2012). User-defined Motion Gestures, *Japan CORE Project Workshop 2012*.

2. Tu, H., Ren, X. and Zhai, S. (2012). A Comparative Evaluation of Finger and Pen Stroke Gestures, *Japan CORE Project Workshop 2012*.

3. Fu, Y., Tu, H. and Ren, X. (2011). Comparison between Ring and Flicking Scrolling Techniques for Document Navigation in Touch-based Mobile Devices, in *Proceedings of FIT2011*, 669-670.

4. Hayashi, Y., Tu, H. and Ren, X. (2011). Comparison between Direct and Indirect Input Techniques on Touch-based Devices, in *2011 Shikoku-section Joint Convention of the Institutes of Electrical and related Engineers*, 334.

5. Tu, H. and Ren, X. (2011). Investigation of Gesture Performance in Different Input Styles, *Proceedings of ISFT 2011: International Symposium on Frontier Technology*, 19-22.