Comparison of Phone-based Distal Pointing Techniques for Point-Select Tasks

Mohit Jain[1], Andy Cockburn[2] and Sriganesh Madhvanath[3]

[1] IBM Research, Bangalore, India
mohitjain@in.ibm.com
[2] University of Canterbury, Christchurch, New-Zealand, [1, 3] HP Labs, Bangalore, India

Abstract. Many different mobile phone-based distal pointing techniques have been proposed and evaluated. Interaction with distant large-screen displays, including interactive TV, requires active pointing and selection of target items. In this paper, we empirically compare four different phone-based distal pointing techniques for point-select tasks. Results show that participants prefer a discrete pointing technique using the phone's touchscreen as a touchpad. This method also achieved the highest accuracy among the techniques studied, with comparable speed. We discuss the implications of our findings for distal interaction.

Keywords: Interactive TV, iTV, Phone, Distal Pointing, Evaluation.

1 Introduction

Interactive TV (iTV) shifts the nature of interaction with distant displays from being predominantly passive consumption of traditional TV to more active control of media and content. Accompanying this shift is an increase in the range of interactive menus displayed on screen and the need for rich input to perform complex tasks including point-and-select. Conventional remote controls consisting of an array of buttons are not ideal for these tasks [11]. This has led to the development of novel distal interaction techniques using phones [4], [9], game controllers [6], hand gestures [17], and novel hardware [5], [10], as input devices. In particular, phones offer several advantages: they are computationally powerful [3], have various built-in sensors, and support wireless communication.

Pointing at targets is a fundamental component of interacting with distant displays. Previous research has successfully used phone's built-in sensors such as *touchscreen, camera, accelerometer, gyroscope,* etc. to sense user's pointing input. *Touch* input, sensed using the touchscreen have been found to work well for distal interaction [7], [12], [13] due to the users' familiarity with laptop touchpads. Applications like Air Mouse [1] allow in-air movements of the phone for interaction. It uses a combination of *accelerometer and gyroscope* data to enable direct pointing, similar to a laser-pointer. The accelerometer measures the acceleration applied to the device, whereas the gyroscope provides the angular orientation of the device in space. The data ob-

[*] Part of this work was done while the first author was an intern at HP Labs India.

tained from these sensors is accordingly mapped to a pointer position on the display. (Techniques tracking phone's camera and flashlight [16] have also been proposed to sense motion inputs, however due to low accuracy and extra hardware requirements, we are not considering them.) Touch and motion are the two most commonly used input techniques, and an understanding of their impact on distal interaction is crucial.

On the distal display, there are two ways to provide visual feedback while selecting an item, irrespective of the input action – (a) *Continuous cursor*, in which a cursor moves continuously over the display (*e.g.*, Gmote [9]), and (b) *Discrete selection-block*, in which a selection-block moves discretely between selectable items (*e.g.*, Apple TV [2]). Note that continuous cursor is target agnostic, while discrete selection-block is target dependent. Continuous and discrete methods converge if targets are a single pixel in size with zero pixels of separation between them. To date, both types of feedback have been widely used, but there has been little work on how on-screen feedback affects distal interaction performance for point-select tasks.

In this paper, we study the relative merits of *continuous* versus *discrete* feedback for pointing interaction, when input is provided through *touch* sensed using the phone's touchscreen and through *movement in the air* sensed using the combination of the phone's gyroscope and accelerometer. We study four distal pointing techniques that satisfy our iTV design goals of single-handed usage with minimal physical and mental load: a) Continuous Touchpad, b) Continuous inAir, c) Discrete Touchpad, and d) Discrete inAir. We conducted a study varying the size and separation of target items to find the relationship between techniques and point-select task difficulty. As these phone-based distal interaction techniques have all been described in prior literature, the contribution of this paper is the empirical comparison of these techniques for point-select tasks. Our study showed that the participants prefer Discrete Touchpad over the other techniques. It also achieved the highest accuracy with comparable speed. We conclude with a discussion of the design implications for distal interaction.

2 Design Goals and Techniques

For phone-based interaction with iTV, we have identified a set of design goals important for end-user adoption: (a) *Single-handed Interaction*: remote controls are usually used with one hand [5]; (b) *Minimal Visual Attention Switching*: the controller should not require direct attention, in order to allow users to focus on the distal display, because attention switching has been found to result in higher error rates [15]; (c) *Minimal Learning Curve*: use of a controller should not require expertise (many pointing techniques described in the literature [4], [5], [13] were found to have a strong learning curve); (d) *Minimal Physical Load*: the controller should use small wrist or thumb movements for interaction to minimize physical load, because techniques requiring arm movement increase fatigue due to the *gorilla-arm-effect* [5].

We studied the following four distal interaction techniques that satisfied our design goals. A *single-tap* on the phone's touchscreen is used for selection.

Continuous Touchpad (CT): In CT, the *touchscreen* of the phone acts as a laptop's touchpad (similar to Gmote [9]), allowing the user to move a cursor over the display

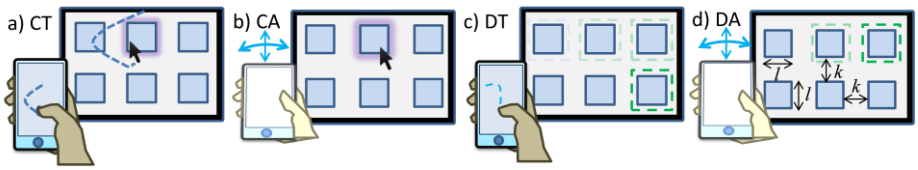

Fig. 1. a-d shows the four interaction techniques. d also shows the definition of item size ('*l*') and gap size ('*k*').

(Fig. 1a). For the mapping between display and touchpad, we used standard Windows XP's cursor acceleration curve, such that a single rapid swipe across the touchscreen allows the cursor to traverse the entire width of the distal display.

Continuous inAir (CA): CA uses the built-in motion sensors of the phone to determine angular orientation and acceleration of the phone and thus control cursor location (similar to Air Mouse [1]). The user moves the phone in air with small wrist-motions to move the cursor (Fig. 1b) to the required target item, followed by a single tap on the phone's touchscreen to trigger selection. The mapping is such that a 30-degree device rotation along the horizontal x-axis, allows the cursor to traverse the width of the display. This mapping value, also known as Control Display gain (CD gain), was chosen using a pilot study with 5 participants. We tested four values of CD gain (20, 25, 30 and 35 degrees) and found maximum accuracy with 30-degrees.

Discrete Touchpad (DT): DT is the same as CT, except that a *discrete selection-block* replaces the cursor as the feedback mechanism, and hence it supports clutching (*i.e.*, unlike a cursor, the selection-block method can never be on a non-selectable or dead space) (Fig. 1c). For example, to move to the selectable item to the immediate right, the user is required to make a small swipe on the touchscreen towards the right. The user can perform a longer swipe to move rapidly through the selectable items.

Discrete inAir (DA): DA is the same as CA, except that a *discrete selection-block* replaces the cursor (Fig. 1d). The user moves the phone in the air with small wrist-motions to move the selection-block. DA has the same CD gain as that of CA.

3 Experiment

3.1 Participants and Apparatus

Twelve male employees of an IT organization participated in the study (mean age 25.6 years, sd 3.4). All but one were right-handed; all reported using computers for 8-10 hours a day; and all but one had previously used a touchscreen phone on a regular basis for 6 months or longer. None had previously interacted with a distant display using a phone. Participants were given a ~$10 voucher as reward.

An iPod Touch 4th generation (display resolution: 960×640) was used for the four interaction techniques. A 42" LCD display (1360×765) connected to a computer as an external monitor was used to simulate the iTV. The display was positioned at eye level and participants sat comfortably on a sofa (with hand-rest) placed 10 feet from the display. The phone and computer interacted over a local wireless network.

3.2 Task, Factors, and Design

Most previous research [4], [11], [14] has used the traditional Fitts' Law point-select task [8] for evaluation. However, the standard Fitts' Law experimental manipulation of controlling Index of Difficulty (the log of the ratio of amplitude to target width) is likely to misrepresent the difficulty of discrete targeting. Thus we performed a variation of the Fitts' Law task. The display was divided into a grid of square-shaped regions of side length ($l + k$) each of which had a square target of side l at its center. k is the gap between two consecutive items (Fig. 1d). The l and k values were varied, which accordingly varied the sizes of the items and the distances between them, respectively. The different values of l and k were selected by studying state-of-the-art iTV interfaces such as Apple TV, and Roku. Nothing was displayed on the phone's screen.

A within-subject design was used for the study. The independent variables were: *Feedback* (2: Continuous, Discrete), *Input* (2: Touchpad, inAir), *Item size l* (3: 100, 60 and 40 pixels), and *Gap between items k* (3: 80, 40 and 20 pixels). The 9 layouts formed by combining the different item size and gap values had different numbers of targets: 28 (# of rows: 4, # of columns: 7; item size: 100, gap: 80), 45 (5, 9; 100, 40), 66 (6, 11; 100, 20), 45 (5, 9; 60, 80), 91 (7, 13; 60, 40), 153 (9, 17; 60, 20), 66 (6, 11; 40, 80), 153 (9, 17; 40, 40), and 264 (12, 22; 40, 20). No item touched the display border, permitting overshooting for continuous techniques, even for the border items.

3.3 Procedure

All the tasks were performed while sitting, holding the phone in the dominant hand (in portrait mode) (Fig. 1). The facilitator demonstrated each of the techniques. Following a practice set of 30 trials with varying target density, participants performed a set of 30 trials for each technique for each of the nine layouts, in a random order. The ordering of the techniques used a Latin Square to mitigate learning and fatigue effects.

In each trial, one of the on-screen items was randomly selected as the target and shown in green. The participant was required to select this item, as accurately and as quickly as possible. When the cursor or selection-block hovered over an item, the item color changed to yellow, providing visual feedback. A correct response is received when the participant triggered selection using a single-tap on phone's touchscreen, while the cursor or selection-block was hovering over the correct target item. Selection of the wrong item was shown by turning it red. Participants had to keep trying until a correct response was received, in order to measure error rate. After each technique, participants were asked to rate the technique in terms of perceived speed, perceived accuracy, and measures from NASA TLX, on a 5-point Likert-scale (for instance, for *fast* 1was very slow and 5 was very fast).

4 Results

The total number of trials = 4 techniques × 12 participants × 3 item sizes × 3 gap sizes × 30 trials = 12960. Participants took ~70 minutes to complete the study, including breaks of 2-5 minutes between techniques.

a)

Tech.	Accuracy	Sel. Time	Errors
CT	93% (7%)	2047 (332)	49% (6%)
CA	81% (7%)	1491 (176)	140% (70%)
DT	96% (2%)	1894 (196)	4% (2%)
DA	86% (5%)	1516 (200)	16% (7%)

b)

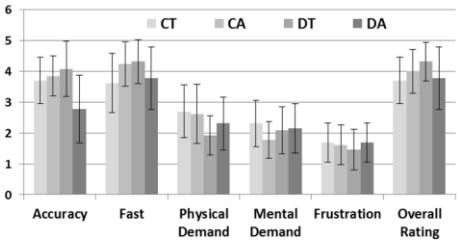

Fig. 2. a) Performance measure means with standard deviation in brackets. b) Likert-scale rating for the techniques (with error bars showing standard deviation)

Our analysis uses a 2×2×3×3 RM-ANOVA for factors *Feedback* (2) × *Input* (2) × *Item size* (3) × *Gap size* (3), and a 4×3 RM-ANOVA for Technique (4) × Distance (small, medium, large). Here distance is the number of items between the starting item and the target item, which can vary from 1 to 20. For analysis, we divided this into three distance levels – small (1-5 items), medium (6-13) and large (14-20).

4.1 Accuracy

Accuracy is defined as the proportion of times the target item was selected correctly on the first attempt. Participants achieved the highest accuracy of 96% with DT (Fig. 2a). The 4×3 ANOVA showed CA was least accurate, with DT more accurate than DA and CA: $F_{3,33}$=23.9, p<0.001. The 2×2×3×3 analysis showed a significant effect for input ($F_{1,11}$=48.5, p<0.001), with touchpad (m=94%, sd=5%) more accurate than inAir techniques (m=83%, sd=7%). For feedback, discrete (m=91%, sd=7%) outperformed continuous (m=87%, sd=9%) techniques, with $F_{1,11}$=8.5, p<0.05.

As expected, there was a significant main effect for item size ($F_{2,22}$=9.2, p<0.01), with larger items being more accurately selected (Fig. 3a). Importantly, there was a significant *feedback* × *item size* ($F_{2,10}$=4.9, p<0.05), and *feedback* × *gap size* ($F_{2,10}$=21.4, p<0.001) interactions. For *continuous* methods, the accuracy decreased with decreasing item size (Fig. 3a), and increased with decreasing gap size (Fig. 3c). This shows that continuous methods are better suited for large items with small gaps between them. The 4×3 analysis found a significant effect for distance ($F_{2,22}$=13.7, p<0.001), and a strong interaction between technique and distance ($F_{6,66}$=7.3, p<0.01). Post-hoc analysis revealed that selecting targets at small distances was significantly more accurate than medium and large distances, with p<0.001 (Fig. 3e). Higher inaccuracy with larger distances may be because of overshooting effects, as found in [4]. This effect was more noticeable for the inAir techniques (Fig. 3e), where the participants reported difficulty in controlling the cursor over longer distances. On analyzing overshooting, we found 67.3% (sd=32.6) of the erroneous selections for DT and DA were due to erroneously selecting items situated a hop away from the target item.

For each selection task, multiple attempts were allowed. Note that this allows error rates to be higher than 100%, as in the case of CA (m=140%). Participants took the fewest attempts with DT, with an error rate of 4% (Fig. 2a). Error rate analysis showed an effect similar to that of accuracy, the details are omitted for brevity.

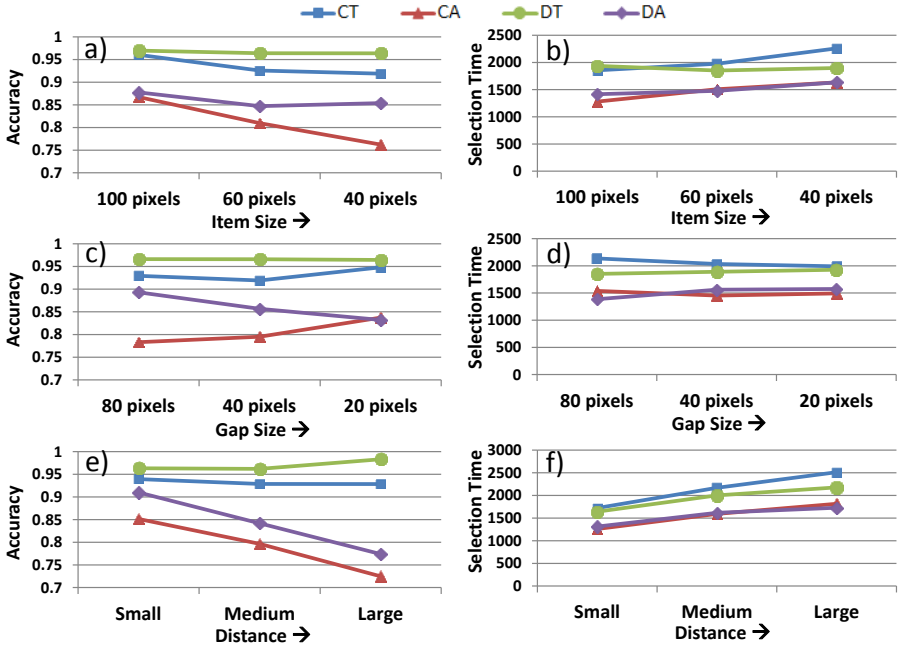

Fig. 3. a) Accuracy and b) Selection Time respectively for different item sizes (a,b), gap sizes (c,d) and distances (e,f)

4.2 Selection Time

Selection Time is the error-free target selection time (trials containing errors were removed). Although CA was least accurate, it was the fastest with an average selection time of 1491 ms showing the speed-accuracy tradeoff. The 4×3 ANOVA showed CA, DA, and DT to be significantly faster than CT ($F_{3,33}$=28.1, p<0.001) (Fig. 2a). The 2×2×3×3 ANOVA showed significant main effects for *input* ($F_{1,11}$=36.3, p<0.001), with *inAir* techniques (m=1504, sd=185) being significantly faster than *touchpad* (m=1971, sd=273). No significant difference was observed between the two feedback methods for selection time.

There was a significant main effect for *item size* ($F_{2,22}$=31.7, p<0.001) with large items being the fastest to select (Fig. 3b). Similarly, large gaps resulted in smaller selection time (Fig. 3d). There were significant *feedback* × *item size* ($F_{2,22}$=16.9, p<0.01) and *feedback* × *gap size* ($F_{2,22}$=17.1, p<0.001) interactions (Fig. 3b, Fig. 3d). For *continuous* methods, the selection time increased with decreasing item size (Fig. 3b), while it remained unaffected by gap size (Fig. 3d). This shows that continuous methods are better suited for large items. For *discrete* methods the selection time remained unaffected by item size, and increased with decreasing gap size; this may be because of the increased density of items. The 4×3 ANOVA found a significant effect for distance ($F_{2,22}$=178.4, p<0.001), and a strong *technique* × *distance* interaction ($F_{6,66}$=6.7, p<0.01). All techniques were significantly faster at selecting targets at small distances, compared to medium and large distances.

4.3 Subjective Assessment and Feedback

Participants' responses to the 5-point Likert scale questions showed that DT received the highest overall rating. DT was also assessed as the fastest, most accurate, involving the least physical effort, and least frustrating (Fig. 2b). All the participants stated that as none of the techniques required looking at the phone screen, it helped them to be more effective. CT was rated the worst in terms of physical and mental load, and for inAir techniques, participants noted some learning curve. Five participants had difficulty with the continuous cursor techniques, citing problems with initially locating the cursor (*e.g.*, *"hard to find the cursor when it has gone to the border of the screen"*) and with making precise selections (*e.g.*, *"selecting smaller targets is tough"*). This may be due to the overshooting effect, resulting in erroneous selections performed while the cursor was in the dead space between target items. On the other hand, participants praised the effectiveness of discrete methods for selecting smaller targets.

Participants liked that the touchpad was *"highly sensitive"*. Three participants complained that moving diagonally towards the upper-left of the touchpad was difficult using the thumb. For inAir techniques, four participants praised the speed offered, while two participants asked to reduce the sensitivity, as it was causing errors while selecting nearby targets. Using CA with small target size, participants complained that *"tapping for selection resulted in slight phone movement, leading to erroneous selection"*. This error was exacerbated due to cursor jitter resulting from hand unsteadiness, as reported in [14]. Hence, for using CA, the application design should have targets of medium to large item size and considerable gap between consecutive items. However all participants *"enjoyed"* using the inAir techniques, which may be due to the novelty factor.

5 Discussion and Design Implications

Overall, the study revealed that touchpad-based techniques were significantly more accurate for point-select tasks. The inAir techniques were perceived as enjoyable, as also reported in [15]. However, participants preferred using touch over motion input, potentially due to accuracy. In terms of feedback, the discrete methods were more accurate than their continuous counterparts, with comparable (or faster) selection times. This is interesting because most current distal pointing systems [9], [10] use continuous feedback. Our results suggest that discrete methods may be preferable unless there is a strong need for a continuous modality, *e.g.*, to support drawing and annotation on screen. Our results also lead us to specific design recommendations for such interactions:

Layout Considerations: For touchpad-based techniques, moving diagonally towards the upper-left of the touchpad was reported to be difficult; hence less-used items should be placed at the upper-left corner region of the display. For continuous feedback, we found overshooting effects, similar to [4]. As a design decision, continuous techniques should always be implemented with gaps between consecutive targets, and performing the selection action while in the dead space should not be an error.

Finger-up Selection: The high error rates with inAir techniques resulted from difficulty in controlling the cursor, specifically from phone movement while tapping for selection. A design solution could be to initiate the inAir movements when a finger

touches the phone's screen, with finger-up used for triggering selection, as that might minimize phone movement for selecting items. Finger-up for selection may be counter-intuitive to the current mental mapping and hence require further research.

Walk-up-and-use System: Interactive systems installed at public locations are walk-up-and-use system, and hence require maximal accuracy with minimal learning curve. For such systems, CT seems like a wise choice, as CA has the highest error rate, while DT and DA requires learning to minimize the overshooting errors.

Given the generic nature of the point-select tasks, our findings and design recommendations are applicable beyond iTV to distal interactions in general. The present study results are based on a limited number of male right-handed participants, with a brief continuous interaction in a lab setting. In the future, we propose a longitudinal study with users from different genders and handedness, and diverse background to explore and evaluate real-world performance of different distal interaction techniques.

References

1. Air Mouse by Movea, https://itunes.apple.com/us/app/air-mouse/id402762122?mt=8
2. Apple Remote app, http://itunes.apple.com/us/app/remote/id284417350?mt=8
3. Ballagas, R., *et al.*: The Smart Phone: A Ubiquitous Input Device. In: IEEE Pervasive Computing, vol. 5, pp. 70--77. IEEE (2006).
4. Boring, S., Jurmu, M., Butz, A.: Scroll, Tilt or Move It: Using Mobile Phones to Continuous Control Pointers on Large Public Displays. In: OzCHI, pp. 161--168. ACM (2009).
5. Choi, S. *et al.*: RemoteTouch: Touch-Screen-like Interaction in the TV Viewing Environment. In: CHI, pp. 393--402. ACM (2011).
6. Cox, D., Wolford, J., Jensen, C., Beardsley, D.: An Evaluation of Game Controllers and Tablets Controllers for Interactive TV Applications. In: ICMI, pp. 181--188. ACM (2012).
7. Enns, N.R.N., MacKenzie, I.S.: Touchpad-Based Remote Control Devices. In: CHI, pp. 229--230. ACM (1998).
8. Fitts, P.M.: The information capacity of the human motor system in controlling amplitude of movement. In: Journal of Experimental Psychology, vol. 47, pp. 381--391 (1954).
9. Gmote, http://www.gmote.org/
10. Loop Pointer, http://www.hillcrestlabs.com/loop/
11. Mackenzie, I.S., Jusoh, S.: An Evaluation of Two Input Devices for Remote Pointing. In: EHCI, pp. 235--249. Springer-Verlag (2001).
12. MacKenzie, I.S., Oniszczak, A.: A Comparison of Three Selection Techniques for Touchpads. In: CHI, pp. 336--343. ACM (1998).
13. McCallum, D.C., Irani, P.: ARC-Pad: Absolute+Relative Cursor Positioning for Large Displays with a Mobile Touchscreen. In: UIST, pp. 153--156. ACM (2009).
14. Myers, B.A., *et al.*: Interacting at a Distance: Measuring the Performance of Laser Pointers and Other Devices. In: CHI, pp. 33--40. ACM (2002).
15. Rashid, U., Kauko, J., Haakila, J., Quigley, A.: Proximal and Distal Selection of Widgets: Designing Distributed UI for Mobile Interaction with Large Display. In: MobileHCI, pp. 495--498. ACM (2011).
16. Shirazi, A.S., Winkler, C., Schmidt, A.: Flashlight Interaction: A study on mobile phone interaction techniques with large displays. In: MobileHCI, pp. 1--2. ACM (2009).
17. Vatavu, R.D.: User-Defined Gestures for Free-Hand TV Control. In: EuroITV, pp. 45--48. ACM (2012).