This is an electronic reprint of the original article.
This reprint may differ from the original in pagination and typographic detail.

Ko, Yu Jung; Putkonen, Aini; Aydin, Ali Selman; Feiz, Shirin; Wang, Yuheng; Ashok, Vikas; Ramakrishnan, IV; Oulasvirta, Antti; Bi, Xiaojun

Modeling Gliding-based Target Selection for Blind Touchscreen Users

Published in:
MobileHCI '21: Proceedings of the 23rd International Conference on Mobile Human-Computer Interactionn

DOI:
[10.1145/3447526.3472022](https://doi.org/10.1145/3447526.3472022)

Published: 27/09/2021

Document Version
Peer reviewed version

Please cite the original version:
Ko, Y. J., Putkonen, A., Aydin, A. S., Feiz, S., Wang, Y., Ashok, V., Ramakrishnan, IV., Oulasvirta, A., & Bi, X. (2021). Modeling Gliding-based Target Selection for Blind Touchscreen Users. In *MobileHCI '21: Proceedings of the 23rd International Conference on Mobile Human-Computer Interactionn: Mobile Apart, Mobile Together* (pp. 1-14). [29] ACM. <https://doi.org/10.1145/3447526.3472022>

Modeling Gliding-based Target Selection for Blind Touchscreen Users

Yu-Jung Ko
yujko@cs.stonybrook.edu
Department of Computer Science,
Stony Brook University
USA

Shirin Feiz
sfeizdisfani@cs.stonybrook.edu
Department of Computer Science,
Stony Brook University
USA

IV Ramakrishnan
ram@cs.stonybrook.edu
Department of Computer Science,
Stony Brook University
USA

Aini Putkonen
aini.putkonen@aalto.fi
Department of Communications and
Networking, Aalto University
Finland

Yuheng Wang
yuhewang@cs.stonybrook.edu
Department of Computer Science,
Stony Brook University
USA

Antti Oulasvirta
antti.oulasvirta@aalto.fi
Department of Communications and
Networking, Aalto University
Finland

Ali Selman Aydin
aaydin@cs.stonybrook.edu
Department of Computer Science,
Stony Brook University
USA

Vikas Ashok
vganjigu@odu.edu
Department of Computer Science,
Old Dominion University
USA

Xiaojun Bi
xiaojun@cs.stonybrook.edu
Department of Computer Science,
Stony Brook University
USA

ABSTRACT

Gliding a finger on touchscreen to reach a target, that is, touch exploration, is a common selection method of blind screen-reader users. This paper investigates their gliding behavior and presents a model for their motor performance. We discovered that the gliding trajectories of blind people are a mixture of two strategies: 1) ballistic movements with iterative corrections relying on non-visual feedback, and 2) multiple sub-movements separated by stops, and concatenated until the target is reached. Based on this finding, we propose the mixture pointing model, a model that relates movement time to distance and width of the target. The model outperforms extant models, improving R^2 from 0.65 for Fitts' law to 0.76, and is superior in cross-validation and information criteria. The model advances understanding of gliding-based target selection and serves as a tool for designing interface layouts for screen-reader based touch exploration.

CCS CONCEPTS

• **Human-centered computing** → **Pointing; Accessibility theory, concepts and paradigms.**

KEYWORDS

gliding-based target selection, pointing, accessibility

ACM Reference Format:

Yu-Jung Ko, Aini Putkonen, Ali Selman Aydin, Shirin Feiz, Yuheng Wang, Vikas Ashok, IV Ramakrishnan, Antti Oulasvirta, and Xiaojun Bi. 2021. Modeling Gliding-based Target Selection for Blind Touchscreen Users. In *23rd International Conference on Mobile Human-Computer Interaction (MobileHCI '21)*, September 27-October 1, 2021, Toulouse & Virtual, France. ACM, New York, NY, USA, 14 pages. <https://doi.org/10.1145/3447526.3472022>

1 INTRODUCTION

Touchscreen devices such as smartphones and tablets have become inextricably tied to the daily lives of people, including those who are either visually impaired or blind. For example, smartphone has become their go-to device for communication and for accessing digital information. Although voice input is becoming increasingly popular, finger touch remains one of the dominant input modalities that they use to interact with touchscreen devices. Voice may raise privacy concerns or is socially inappropriate in public environments such as in a train or library. In these scenarios, finger touch is the main input modality for blind people.

One common touch-based interaction blind people perform is gliding-based target selection (a.k.a touch exploration): to reach the target, a blind user glides the input finger on the screen to explore and locate the target while a screen reader speaks the labels of the UI elements under the finger. Such interaction is often used by blind users to explore and learn a new user interface (UI), as well as to select elements when their locations are well known (e.g., during web browsing). In practice, both Android's TalkBack [42] and Apple's VoiceOver [2] screen readers support gliding for reaching a target (Figure 1), and many blind users rely on gliding to type on soft keyboards, as described in previous work such as [3, 8]. Gliding-based selection is also an important technique in many assistive technologies such as Slide rule [28] and Access Overlays [29].

The efficiency of gliding-based interaction is sensitive to the spatial layout of the interface. As a user glides a finger on a touchscreen,

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

MobileHCI '21, September 27-October 1, 2021, Toulouse & Virtual, France

© 2021 Association for Computing Machinery.

ACM ISBN 978-1-4503-8328-8/21/09...\$15.00

<https://doi.org/10.1145/3447526.3472022>

Gliding-based Target Selection

Figure 1: Gliding-based target selection. The user glides her finger on screen to reach the target.

both the locations and sizes of UI elements influence the performance of gliding-based target selection. However, most existing interfaces are designed for sighted users by default, thereby poorly supporting gliding-based interaction for screen-reader users. For example, many interface elements are too small (e.g., less than 12 mm in width), making them error-prone for gliding-based selection; some UI elements are placed far apart for reasons of visual aesthetics, which result in unnecessarily long selection times. There is therefore a need for guidance to re-arrange and re-size touchscreen UI elements to improve the efficiency and accuracy of gliding-based interaction for blind screen-reader users.

This paper contributes to mathematical models of user performance, a powerful tool for capturing tendencies in empirical data in a format that can guide practical design decisions and serve as the basis for computational methods, such as adaptive interfaces and optimization. The model we explore here builds on the tradition of Fitts' law based pointing models. It relates the target selection time (MT) to the properties of the target, in particular its width and distance. Previously such models have been developed almost exclusively for sighted users and graphical interfaces, limiting their use in design for blind users. The model proposed here allows a designer to estimate the average accessing time to UI elements of a particular design at an early stage, ruling out inferior options. While existing target selection models are based on Fitts' law [15, 16, 33], and assume that visual feedback is present, our model captures motor phenomena specific to finger gliding with no visual feedback.

In this paper we propose the *mixture pointing model* (Equation 1), which relates target selection time MT with the finger travel distance A and target width W in gliding-based target selection tasks for blind people:

$$MT = a + bA + c \log_2\left(\frac{A}{W} + 1\right), \quad (1)$$

where a , b , and c are empirically determined parameters. We propose the model based on the observation that blind people adopt a mixed strategy in gliding-based target selection: some actions follow the iterative corrections model which is nonetheless captured by Fitts' law [15, 16, 33], while others follow a previously unreported strategy where multiple sub-movements separated by stops

are "concatenated" until the target is reached. These movements are understandable when considering the need for auditory feedback for localization while using screen-readers, which requires the user to wait for output. The model reflects the mixture of these two strategies: the term $c \log_2(\frac{A}{W} + 1)$ captures the iterative-correction strategy (a.k.a the Fitts' law component), and bA captures the multi-stop movement.

The empirical study we report here suggests that this model outperforms Fitts' law. It increases R^2 from 0.65 (Fitts' law) to 0.76, and has higher prediction accuracy as measured by cross-validation and information criteria metrics which control for overfitting. This mixture model also outperformed a number of previously reported pointing models [13, 19, 34, 35], measured by R^2 , cross-validation, and information criteria.

The mixture pointing model advances our understanding of how blind people select targets on a touchscreen, and suggests new implications for designing accessible touchscreen interfaces. In contrast to Fitts' law, which specifies that both finger travel distance A and target width W are equally important in determining MT , the model suggests that A plays a more important role than W in determining MT because of the multi-stop component bA , thereby indicating that shortening the finger travel distance (A) is more critical than enlarging the target width (W) for blind people. The new model can serve as a tool to assist accessible user interface design, as demonstrated in an example in which a menu design is optimized for blind people.

2 RELATED WORK

We first discuss related work on touchscreen accessibility, and then review previous research about pointing models, motor actuation of blind people, and applications of pointing models.

2.1 Touchscreen Accessibility

Blind people rely on assistive technologies, such as VoiceOver in Apple [2] and TalkBack in Android [42], to interact with touchscreen devices. Both of them are screen reader services which speak the labels of UI elements as users glide their finger over them on the screen. Besides gliding-based target selection, blind users may repetitively swipe to sequentially access targets. The downside of swiping is that it may be time-consuming especially for searching through a long list. Voice input is becoming increasingly popular among blind people thanks to the advances in speech recognition technology [23]. However, speech input may raise privacy concerns and is socially inappropriate in public places such as in public transport, library, or meeting.

Previous research has explored various techniques to improve the accessibility of touchscreen devices. Slide rule [27, 28] is a well defined multi-touch action that allows users to access application such as web browsing by using auditory guidance. Kane et al. [29] extends the accessibility of touchscreens by applying overlays upon the screens and provide guidance to help the users pinpoint the target location. Beyond ordinal touchscreens surfaces, Guerreiro et al. [21] explores the accessibility of multi-touch interaction of the blind users with large touchscreens. Among various accessibility techniques, gliding-based pointing remains a significant role, especially for exploring unknown UIs and soft keyboard typing for

people with visual impairments. Next, we discuss work related to understanding and modeling pointing behaviors.

2.2 Pointing models with visual feedback

As one of the classical topics in HCI, pointing performance has been widely studied for more than half a century. One of the earliest and most prevalent pointing models is Fitts' law [15, 16, 20, 32, 33, 33, 34], which suggests a logarithmic relationship between target selection time (MT) and the relative spatial precision of a task $\frac{A}{W}$. Paul Fitts originally formulated this logarithmic relationship by assuming that a human behaves like an information-transmission channel of limited capacity, and in a rapid aimed movement, the amount of information to be transmitted was $\log_2(\frac{2A}{W})$ (a.k.a. the index of difficulty ID). Assuming that the transmission rate is constant, this results in a linear relationship between MT and ID . Since Fitts' law was proposed in 1954, abundant research [44, 45] has been conducted to refine and re-interpret ID . The most widely adopted variant of ID shows better correlation with experimental data [32]: in the HCI community $ID = \log_2(\frac{A}{W} + 1)$. We too adopt this formulation of ID in this paper.

Although Fitts' law was originally formulated from an information-theoretic basis, it can be explained by motor-control or manual-control theories. One such explanation [12, 30] is the iterative-correction model, which states that the pointing movement consists of discrete sub-movements guided by visual feedback, and each sub-movement covers a fixed proportion of remaining distance to the center of the target. Another explanation is the *stochastic optimized sub-movement* model [34, 35], which states that an aimed movement involves a primary sub-movement and optional secondary corrective sub-movements. Additionally, studies on the dynamics of rapid aimed movement (e.g., [36]) have revealed that the pointing behavior with visual feedback consists of an initial large, often ballistic movement, followed by multiple gradual correction movements. Overshooting and iterative corrections are also observed. Additionally, research has also been carried out to model ballistic movement, a rapid voluntary movement in which the visual feedback for path correction is not possible. For ballistic movements, the movement time was found to be related to the square root of the movement amplitude [19].

It is unknown whether these previously proposed models can be generalized to blind users. We included multiple of them in model selection, including Fitts' law, ballistic movement model [19], and stochastic optimized submovement models [34, 35]. Our evaluation showed that the mixture model proposed in this work outperformed all other candidates in model fitness performance.

2.3 Pointing Performance of Blind People

In case of blind users, the lack of visual feedback is partially compensated by other sensory modalities [4, 38]. For example, a study [39] showed that blind users can interact effectively with others without visual experience. A wearable device proposed in [22] utilizes haptic cues to help users with visual impairments to fetch objects. In another work [26], haptic feedback provided on wrists was shown to be effective in improving speed and accuracy of target finding tasks on a 2D surface for blind users. On the other hand, as visual feedback is crucial in guiding interaction, the lack of visual signals

inevitably affects interaction efficiency. The following research is especially related to modeling the pointing behavior of blind users.

An experiment [14] on blindfolded and blind users showed that Fitts' law can be applied to blind users on a touchscreen with vibrotactile feedback. With the approximate locations of the targets in their minds, blind users can effectively acquire targets by moving hands over the air. On the other hand, a study [13] on cursor-based target acquisition tasks on a linear or grid menu showed that the time for selecting a target is linearly related to the distance, which was defined as the number of menu items the cursor needs to travel. We include this model as a candidate in our model selection procedure.

Compared to previous research, a major difference in our work is that we studied the pointing action in a new setting: gliding fingers on a touchscreen with audio feedback to select targets. It differs from the previous study [14] where blind users move their hands over the air to select targets, and is distinct from the previous study [13] which used a mouse as the input device.

2.4 Applications of Pointing Models

Mathematical models are formal and simplified representations of the real world, and have served as foundations in a number of applied fields. In HCI, Fitts' law and its variants have been used to compare the performance of input devices as well as the basis of the computational design of layouts [37], including keyboards and graphical layouts and menus. It has informed the design of better interaction techniques for e.g., menus [5] and soft keyboards [6]. Pointing models for people with special needs would serve as foundations for ability-based interface design and optimization. For example, SUPPLE [17, 18] and associated related work optimized user interfaces based on models that can predict the performance of people with special needs. Recent research created models for older adults and used it to optimize mobile interfaces (e.g., soft keyboard layout) to suit older adults' interaction abilities [40]. The presented research contributes a model that predicts blind people's gliding-based pointing performance, which would serve as a tool and building block for accessible user interface design and optimization.

3 EXPERIMENT: UNDERSTANDING GLIDING-BASED TARGET SELECTION

We carried out a controlled experiment to understand the basic gliding-based target selection behavior of blind people. Similar to previous research of investigating pointing performance [33], we followed a canonical 1D target selection task (Figure 2): Blind participants acquired vertical bars on a touchscreen by gliding the finger with the assistance of a screen reader – an assistive technology that reads out the content underneath the input finger. The finger stays in contact with the screen throughout. We manipulated the distance and width of targets.

This task is one of the simplest target selection tasks but still captures the fundamental components of the pointing behavior. Our experiment used screen readers to guide the selection because screen reader services such as VoiceOver on iOS and Talkback on Android have been the *de facto* standard assistive tools for blind users to interact with smartphones and tablets. Such software is

ID	Age	Gender	Input Finger for Touch Interaction	Legally Blind?	Diagnosis	Screen Reader User?
P1	37	Male	Index	Yes	Optic atrophy, Retinitis pigmentosa	Yes
P2	39	Male	Index	Yes	Congenital cataracts	Yes
P3	31	Female	Index	Yes	Surgery mistake	Yes
P4	58	Male	Index	Yes	Detached retina	Yes
P5	54	Male	Index	Yes	Detached retina	Yes
P6	32	Female	Index	Yes	Optic nerve atrophy	Yes
P7	46	Female	Index	Yes	Leber's congenital amaurosis	Yes
P8	36	Female	Index	Yes	Glaucoma	Yes
P9	55	Male	Index	Yes	NA	Yes
P10	45	Male	Index	Yes	Diabetic Retinopathy	Yes
P11	31	Female	Index	Yes	NA	Yes
P12	58	Female	Index	Yes	Nerve Atrophy	Yes
P13	58	Male	Index	Yes	Congenital Cataracts	Yes

Table 1: Demographics of participants.

offered for free to users on major mobile computing platforms, including iOS and Android, and they have been widely adopted.

3.1 Participants and Apparatus

We recruited 13 legally blind participants (6 females) aged between 31 and 58 years (mean = 44.6, std = 10.9) for our IRB-approved study (Table 1). All participants participated under informed consent and were from a non-profit organization that provided rehabilitation services for people with vision impairments. All participants had experience with touch-based devices - tablets and smartphones. They were instructed to perform the one-dimensional target acquisition task as quickly and as accurately as possible on a Google Pixel C tablet, with a screen resolution of 2560x1800 and 308 ppi, which means 1mm on the screen contains around 12.126 pixels.

Figure 2: Left: a blind participant in study. Right: screenshot of the task. The participants glided the input finger from a starting position (blue bar) to select the target (red bar) with the assistance of a screen reader.

3.2 Design

We adopted a within-subject design for a target acquisition task in which participants acquired vertical bars on a tablet screen. The independent variables were the following:

- Finger Travel Distance (a.k.a amplitude A) with 8 levels (in cm): 4.0, 4.8, 5.6, 6.4, 7.2, 8.0, 8.8, and 9.6.

- Target Width (W) with 8 levels (in cm): 0.6, 1.2, 1.8, 2.4, 3.0, 3.6, 4.2, and 4.8.

This independent variables setting is to make sure most combinations of touch interface design are included in the real gliding-based selection. The experiment settings covered the target widths of range 0.6-4.8cm from small to large targets. The distance ranges from 4.0 to 9.6cm which covers the common finger travel distances on touchscreen devices. In total, there were 64 conditions ($8A \times 8W$). For each (A, W) condition, the participants were instructed to perform 6 successful target selection trials described next.

3.3 Procedure

We designed a target acquisition task in which blind users selected targets by gliding their input fingers on a touchscreen to reach the target. At the beginning of each trial (Figure 2), one vertical light blue bar, which represented the starting region, and one vertical light red bar, which indicated the target, were displayed on the screen. To start a trial, the participant first placed their finger on the screen, and then glided the finger over the screen to reach the starting region. The input finger should be in contact with the screen all the time. Voice assistance was provided for this gliding activity in the form of continuous auditory feedback ("Source on the left" or "Source on the right") from the screen reader. As soon as the input finger reached the starting region for the task, the screen-reader indicated the start of the trial by narrating "Task starts", and then notifying the participant about the location of the target by narrating either "Target on the left" or "Target on the right". The participant then moved their finger on the screen to reach the target. Once the input finger reached the target, a notification "Reached Target" was voiced by the screen reader. If the input finger overshot, i.e., went beyond the target, the screen-reader would say "Passed the Target". The participant then had to rectify their finger's location by moving it back to the target. If the participant was able to move their finger back to the target, the screen reader would again say "Reached target". If the participant lifted their finger up outside the target, the selection was considered as a failure. In this case, the screen-reader would say "Target not selected", and the

participant had to repeat the trial. If the participant lifted their finger up from a position within the target, the selection was considered as a success, and the screen-reader would then say "Target selected". The participant then proceeded to take the next trial in which the locations of starting region and the target were changed. The ordering of the conditions was randomized for each participant. The duration of audio feedback ranged between 800 - 1900 ms, summarized in Figure 3.

Figure 3: The Screen-reader provided audio guidance to the user while she moves her finger from the starting bar to the target bar. The duration of audio feedback ranged from 800 - 1900 ms, as shown in Table 2

Message	Duration (ms)
Reached Target	832
Passed Target	877
Target Selected	1117
Task Starts, Target on the left	1903
Task Starts, Target on the right	1883

Table 2: Durations of Audio Feedback Phrases

The auditory cues provided the necessary guidance regarding the gliding direction and feedback on whether the target was selected, which was the crucial information in gliding-based target selection task. Our auditory feedback setting was different from a typical menu design where the label of the menu underneath the input finger would be narrated. We adopted a slightly different design because our study had only one target and the user would have no idea where the target was at the beginning of the trial if the application provided auditory feedback only when the finger reached it. It was a simplification of a realistic gliding behavior. Except that the gliding direction was provided at the beginning of the trial, the rest of the gliding behavior was the same as how a blind user would glide the finger to select an object on a touchscreen.

We collected 5551 trials in total, among which 4992 were successful trials. There were 212 trials (4%) where participants lifted their finger up before even reaching the starting region. These trials did not get started and were removed as outliers. We focused on analyzing the remaining 5339 trials.

3.4 Results

We first examined the movement time and error rates. Movement time was defined as the elapsed duration between the moment the screen-reader indicated "Task Starts" until the moment the finger was lifted up. Error rate was computed as the percentage of failed trials in which the participants lifted their finger up outside the target.

Figure 4: Mean (SD) Movement time by target width W and distance A .

Figure 5: Mean (SD) Error rate by target width W and distance A .

Figure 6: Mean (SD) of number of overshootings by target width W and distance A .

Figures 4, 5 present the mean (SD) movement times and error rates with respect to finger travel distance and target width. For movement times, a repeated measure ANOVA test showed that both width W ($F_{7,84}=17.82$, $p < 0.0001$) and distance A ($F_{7,84}=15.25$, $p < 0.0001$) had a statistically significant effect. The interaction effect of width and distance was also significant ($F_{49,588}=1.388$, $p < 0.05$). For error rates, a repeated measure ANOVA test showed that width W had a significant effect ($F_{7,84}=15.33$, $p < 0.0001$), but not distance A

($F_{7,84}=1.122, p = 0.357$). The interaction effect of width and distance was not significant ($F_{49,588}=1.195, p = 0.177$).

In addition to movement times and error rates, we also analyzed the number of overshootings. An overshooting is the case where the participant did not stop moving their finger after reaching the target. Figure 6 shows the number of overshootings with respect to A and W . It can be observed from Figure 6 that the number of overshootings was relatively larger when the target width was smaller than or equal to 12mm. For target widths larger than or equal to 18mm, the number of overshootings was comparatively smaller. A repeated measure ANOVA test showed that both width W ($F_{7,84}=40.63, p < 0.0001$) and distance A ($F_{7,84}=3.38, p < 0.005$) had statistically significant main effects. The ANOVA test did not show a significant interaction effect for width $\times$ distance ($F_{49,588}=1.017, p = 0.445$).

4 MOVEMENT DYNAMICS

To understand how blind users select targets with gliding, we investigated the movement dynamics, including speed and acceleration of finger movement during the target selection procedure. As touch position data often involved noise and uncertainty, prior to the analysis we applied a Savitzky–Golay filter [41] to smooth the position time-series data and derive the speed and acceleration of each trajectory. Since the sampling of our device was nonuniform, the Savitzky–Golay filter was a reasonable choice in our signal smoothing of position. Instead of using the finite difference of the positions, we adapted the first derivative of the filtered trajectory as the speed and the second derivative as the acceleration. The Savitzky–Golay filter we used has window size = 11 and polynomial order = 2.

Figure 7: Idealized phase space plots (i.e., speed vs. position) for (a) iterative-correction and (b) multi-stop movement strategies. Supposing the user moves their finger from the start region (blue bar) to select the target (red bar), Figure (a) shows how speed varies as finger position changes for iterative-correction movement. It starts with a ballistic movement followed by iterative correction, which is similar to dynamics captured by Fitts’ law as shown in [34]. Figure (b) shows how speed varies as finger position changes for multi-stop movement, which is a concatenation of multiple sub-movements separated by stops until the finger reaches the target.

4.0.1 Two Target Selection Strategies. Examining the movement dynamics, we discovered two types of target selection strategies, as illustrated in (Figure 7).

- **Iterative-Correction Movement.** As shown in Figure 7(a), the user performed a ballistic movement and iteratively corrected it until it reached the target. The finger movement never stopped until the finger lifted up. This type of dynamic was similar to dynamics observed for sighted users in target selection tasks whose task completion time (MT) can be modeled by Fitts’ law, as reported in the literature (e.g., [34, 36]). Blind users were able to perform such a movement probably because they were able to approximately estimate the location of the target after successfully completing a few trials in each condition. Their estimation of the target location could guide them to perform a ballistic movement at the beginning of the trial which would place the finger location near the target; the user can then iteratively correct the finger location until it reached the target.
- **Multi-Stop Movement.** As shown in Figure 7(b), the movement was a concatenation of multiple sub-movements until the finger reached the target. Sub-movements were separated by finger stops.

We classified a trial as iterative-correction or multi-stop movement based on the number of *stops* in the movement. A stop position is a finger position along the trajectory at which (1) the input speed of the finger is low, below 1 cm/s, and (2) its speed is the lower than its immediate “neighbors” (i.e., the preceding and following points). The first criterion requires the finger to nearly stop, and the second criterion ensures at most one stop will be counted when the finger speed monotonically decreases to reach a stop. If the finger trajectory contains more than 1 stop before reaching the target, it is a multi-stop movement, otherwise it is an iterative-correction movement.

Figure 8 shows several examples of the iterative correction and multi-stop movements performed by the participants in our study. For iterative-correction movement, outcomes like Figure 8 (a) contain some brief and mild deceleration followed by acceleration for correction. Therefore, they have a short and flat segment in the middle, but overall they still follow the speed profile illustrated in Figure 7 (a). That is, an acceleration phase followed by a deceleration phase. Figure 9 shows all trials in one particular condition ($A = 0.8cm, W = 1.2cm$): both strategies were present in pointing.

Figure 10 shows the percentages of two strategies across trials within a condition, and across participants. As shown in Figure 10, the percentages of two types of movements were similar across the trials within a condition, but differed across participants. For example, P1 performed the majority of trials with iterative-correction movements, while P12 mostly chose the multi-stop strategy.

5 MIXTURE MODEL FOR TARGET SELECTION

Based on the observation that blind people adopted two strategies for target selection, namely multi-stop movement and iterative-correction movement, we proposed modeling the target selection time with a mixture model to reflect these two strategies separately. In this section, we first propose models for multi-stop movements

Iterative-Correction Movement

Multi-Stop Movement

Figure 8: Examples of iterative-correction movements (a, b and c) and multi-stop movements (d, e, and f). a, b, d, and e are individual trials, and c and f are aggregated trails in one (A, W) condition.

Figure 9: Visualizations of movement dynamics for all trials in the condition where $A = 8.0\text{cm}$ and $W = 1.2\text{cm}$ (as an example). Light blue curves represent iterative-correction strategies and orange curves represent multi-stop strategies. On the left, a time series plot shows the changes in the finger touch position over time. In the middle, a phase space plot shows finger speed at a position. On the right, a Hooke plot shows acceleration over position.

Figure 10: Average percentages of iterative-correction and multi-stop movements by the orders of 6 trials in the corresponding (A, W) condition (up), and by participant (bottom).

and iterative-correction movements, respectively, and then combine them to form a mixture model.

5.1 Model for Multi-Stop Movement

This model assumes that the target selection movement consists of multiple sub-movements separated by finger speed indicating a stop. Therefore, the overall movement time (MT_{ms}) is calculated as:

$$MT_{ms} = \sum_{i=1}^n MT_i, \quad (2)$$

where MT_i is the duration of each sub-movement and n is the total number of stop-separated sub-movements.

Without more information about sub-movement MT_i , we assume each sub-movement travels a constant distance τ , and takes a constant time t . Under this assumption, the total number of sub-movements $n = \frac{A}{\tau}$ and the total movement time can be calculated as:

$$MT_{ms} = \sum_{i=1}^n MT_i = n \cdot t = \frac{A}{\tau} \cdot t. \quad (3)$$

Because τ , and t are both constants, Equation 3 shows $MT \propto A$. Assuming there is initiation time for starting the movement, we predict the movement time MT as:

$$MT_{ms} = a_0 + b_0 A, \quad (4)$$

where A is the total travel distance, and a_0 and b_0 are empirically determined parameters.

5.2 Model for Iterative-Correction Movement

The movement pattern in the iterative corrective movement follows Fitts' law. It assumes that the user moved the finger to a position near the target in a ballistic movement, and then iteratively adjusted

the final location to select the target. We proposed to model such movement following Fitts' law [15, 33]:

$$MT_{it} = a_1 + b_1 \cdot \log_2\left(\frac{A}{W} + 1\right), \quad (5)$$

where a_1 and b_1 are empirically determined parameters.

5.3 Mixture Pointing Model

Assuming the observed target selection movements are a mixture of multi-stop and iterative-correction movements where $\alpha (0 \leq \alpha \leq 1)$ proportion of the trials are multi-stop movement and $(1 - \alpha)$ proportion of trials follow iterative correction movements, we can model the total movement time as:

$$MT = \alpha \cdot MT_{ms} + (1 - \alpha) MT_{it} \quad (6)$$

$$= \alpha(a_0 + b_0 A) + (1 - \alpha)(a_1 + b_1 \cdot \log_2\left(\frac{A}{W} + 1\right)) \quad (7)$$

$$= (\alpha a_0 + a_1 - \alpha a_1) + (\alpha b_0) A + (b_1 - \alpha b_1) \cdot \log_2\left(\frac{A}{W} + 1\right). \quad (8)$$

Because α, a_0, b_0, a_1 , and b_1 are all constants, we can simplify Equation 8 to:

$$MT = a + bA + c \log_2\left(\frac{A}{W} + 1\right), \quad (9)$$

where a, b , and c are empirically determined parameters. This mixture pointing model (Equation 9) reflects the mixture of iterative-correction movement ($c \log_2(\frac{A}{W} + 1)$) and multi-stop movement (bA). The coefficients b and c reflect the relative contribution of iterative-correction and multi-stop components to determine MT .

5.4 Model Evaluation

We evaluated the following 7 model candidates with the data collected in our experiments. These model candidates included our newly derived mixture pointing model (Equation 9), Fitts' law, and other previously reported models in literature. We fitted the models,

calculated the information criteria. Besides, we practiced leave-one-group-out to demonstrate the proposed models was not overfitted.

- Mixture pointing model. This is Equation 9, assuming blind users adopted a mixed strategy for selecting targets.
- Fitts' Law. We used the Shannon version [32, 33]:

$$MT = a + b \log_2\left(\frac{A}{W} + 1\right), \quad (10)$$

where a and b are empirically determined parameters.

- Linear Model. This is identical to the multi-stop movement model (Equation 4), which is another sub-component of the mixture model:

$$MT = a + bA, \quad (11)$$

where A is the movement amplitude.

- $\sqrt{A}$ model. This is a ballistic movement model proposed in the literature [19]. Ballistic movement is defined as a rapid voluntary movement in which visual feedback for path correction is not possible. Since it models such movement without visual feedback, we included it in the evaluation. The ballistic movement model is formulated as follows:

$$MT = a + b\sqrt{A}, \quad (12)$$

where A is the movement amplitude.

- $\sqrt{\frac{A}{W}}$ Model. This is the stochastic optimized sub-movement model proposed by Meyer et al. [34, 35]. It assumes that the movement consists of a primary sub-movement and optional secondary corrective sub-movements, and the sub-movements are programmed to minimize MT and keeping low error rates. This model is expressed as follows:

$$MT = a + b\sqrt{\frac{A}{W}}. \quad (13)$$

- $\frac{A}{W}$ Model. This model is similar to the one proposed by Edwards [13], which states that when blind users selects an item in a linear menu or grid menu, the movement time MT is linearly related to distance. The distance is defined as the position of the item in a menu layout. We adapted it to suit our experimental setting by measuring the distance as the number of W units:

$$MT = a + b\frac{A}{W}. \quad (14)$$

- $\left(\frac{A}{W}\right)^{\frac{1}{n}}$ Model. This is the generalized stochastic optimized sub-movement model [35] which defines MT as a function of not only A and W , but also the number of sub-movements:

$$MT = a + b\left(\frac{A}{W}\right)^{\frac{1}{n}} \quad (15)$$

5.4.1 Model Fitness. Using a least-squares fit method, we estimated the parameters of 7 model candidates on 64 mean MT in our experiment (because our experiment had 64 (A, W) pairs). The results are shown in Figure 11. As it can be seen at Figure 11, the mixture pointing model achieved the best performance among all the model candidates. R^2 of fitting result of mixture pointing model is 0.76; Fitts' law has $R^2 = 0.65$. Both linear and $\sqrt{A}$ models performed poorly, with $R^2 < 0.4$.

5.4.2 Information Criteria. The Akaike (AIC) [1] and Watanabe-Akaike ($WAIC$) [43] information criteria were used to examine the model fitness. These criteria, which penalize the number of parameters in case of overfitting, demonstrated the performance of the mixture pointing model over Fitts law and other models. In general, the smaller the information criterion, the better the model is. Figure 11 shows that the mixture pointing model had the lowest values of information criteria ($AIC = 30.06$ and $WAIC = 29.86$), indicating that the mixture pointing model performed the best among all the candidates.

5.4.3 Cross Validation with RMSE. We performed leave-one-user-out validation on the mean movement time. We fitted the model in every other 12 participants and validate with the data from the remaining user. In the validation process, we calculated root mean square error (RMSE). Averages (with standard deviation) of 13 sets of metrics over 7 models above are shown in Fig 11. We found that the mixture pointing model still has the lowest RMSE = 1.11, followed by sub-optimal power model, Fitt' law and $\sqrt{\frac{A}{W}}$ model's RMSE = 1.13.

We also performed leave-one- (A, W) -out cross validation on the mean movement time. Each width and distance pair would be excluded in every iteration as validation data. There are 64 RMSEs with different validation (A, W) pairs of each model and the average of it is reported. The mixture pointing model has lowest RMSE = 0.23, followed by Fitts' Law, $\sqrt{\frac{A}{W}}$ model, and $\left(\frac{A}{W}\right)^{\frac{1}{n}}$ model with RMSE = 0.27. With both leave-one-out cross validation, the mixture model yielded the best result with RMSE.

5.4.4 Model Fitness for Iterative-Correction and Multi-Stop Trials. To further understand the performance of the mixture model, linear model, and Fitts' law model, we analyzed their fitness on iterative-correction and multi-stop trials separately, which were defined according to the criteria outlined in Section 4.0.1. For the iterative-correction trials, the R^2 value of the mixture model was 0.70, and the parameters were $a = 1.71$, $b = 0.08$, and $c = 0.54$. For the multi-stop trials, the R^2 value of the mixture model was 0.57 and the parameters were $a = 2.35$, $b = 0.14$, and $c = 0.42$. The main difference is that on the multi-stop trials, the coefficient of the linear term b ($b = 0.14$) was much greater than the coefficient of the linear term ($b = 0.08$) on iterative-correction trials, suggesting that the linear term ($b \cdot A$) contributed more to MT for the multi-stop trials than for iterative-correction trials. The mixture model also outperformed the linear model and Fitts' law model in each of the two types of trials. The R^2 values of Fitts' law were 0.64 for iterative-correction trials and 0.43 for multi-stop trials, and the R^2 values of the linear model were 0.26 for iterative-correction trials and 0.34 for the multi-stop trials.

These results showed that both the logarithmic term ($c \log_2(\frac{A}{W} + 1)$) and the linear term ($b \cdot A$) contributes to MT in both types of trials; the contribution of linear term ($b \cdot A$) is more prominent for the multi-stop trials than for the iterative-correction trials. The higher value of b in multi-stop trials is expected as described in Section 4.0.1. Neither the Fitts' law nor linear model performed well for each type of data. It is probably because the criteria for

Figure 11: Model parameters and evaluation results of 7 model candidates. Information criteria (AIC and WAIC) measure model fitness after controlling for overfitting (Smaller the value, better the model). As it can be seen from the figure, the mixture model performed best, measured by R^2 , leave-one-user-out RMSE, leave-one-(A, W)-out RMSE, and information criteria.

Mixture Model	R^2	RMSE (user)	RMSE (A, W)	AIC	WAIC
#1	0.75	1.12	0.23	30.92	30.8
#2	0.77	1.11	0.22	27.02	27
#3	0.75	1.12	0.23	30.24	30.08

Table 3: Evaluation results of 3 mixture model candidates. R^2 , leave-one-user-out RMSE, leave-one-(A,W)-out RMSE, and information criteria

classifying the trials into two types of movement are heuristic-based and obtained by examining the dynamics of finger movement. They only approximately separated trails into iterative-correction and multi-stop movements. As shown in Figure 10 (f), there were some trials classified as multi-stop strategies in which a user firstly performed one sub-movement to cover a large distance, followed by small adjustments. On the other hand, as shown in Figure 10(c), the speeds of some trials classified as iterative movement went up and down multiple times before the finger reached the target, indicating that there were a certain degree of multi-stop like movement in this type of trials too.

5.4.5 Other Mixture Model Candidates. We also examined the fitness with other mixture model candidates, by mixing a model that predicts MT with A and a model that predicts MT with $\frac{A}{W}$. More specifically, we examined the following mixture models:

- Model #1. A mixture of $\sqrt{A}$ model proposed in [19] and Fitts' law model:

$$MT = a + b\sqrt{A} + c \log_2\left(\frac{A}{W} + 1\right).$$

- Model #2. A mixture of linear model and $\sqrt{\frac{A}{W}}$ model proposed in [34, 35]:

$$MT = a + bA + c\sqrt{\frac{A}{W}}.$$

- Model # 3. A mixture of $\sqrt{A}$ model and $(\frac{A}{W})^{\frac{1}{n}}$ model which is a generalized approach mentioned in [35]:

$$MT = a + b\sqrt{A} + c \log_2\left(\frac{A}{W} + 1\right).$$

Table 3 shows the evaluation results of three potential mixture model candidates. The results led to two findings. First, all these models outperformed using a single model to predict MT , compared with the performance reported in Figure 11. It indicated that a using a mixed-model strategies was more appropriate than using a single-model strategy to model blind users' gliding behavior. Second, all these models (models #1 - #3) had similar performance, compared with the proposed mixture model (Equation 9): they have similar R^2 , and RMSE, and the differences in information criteria (i.e., AIC, WAIC) are all less than 4 [11]. It indicates these models (models #1, #2, and #3) can all serve as appropriate candidates for modeling blind users' gliding behavior.

6 DISCUSSION

Our main finding is that the movement strategies of people who are blind in gliding-based target selection exhibit a mixed pattern with two main styles. This contrasts the assumptions of Fitts' law. Take the velocity dynamics as an example. Fitts' law predicts that the velocity of rapid aimed movement forms a bell-shaped curve, reaching a peak near the middle of the movement and gradually decreasing to stop at the end. While some of the selection trials followed this pattern, a large proportion of trials followed a different dynamics: a concatenation of multiple sub-movements separated by stops. A plausible explanation for this strategy is that due to the lack of visual feedback, blind people search for the target position by performing multiple sub-movements.

Based on this finding, we proposed a mixture pointing model, which is a mixture of Fitts' law model and multi-stop movement model, to predict blind users' target selection time. The results showed that this model outperforms Fitts' law, measured by R^2 and information criteria which control for overfitting.

6.1 Design implications

The mixture pointing model has implications on accessible user interface design. First, the model (Equation 9) shows that the movement amplitude A plays a more important role than W in determining MT , because of the added multi-stop term $b \cdot A$. In other words, targets placed closer to the point of origin will be selected quicker; an effect that is underestimated by Fitts' law. This result has implications in cases where items must be prioritized: our recommendation is to place the items to be prioritized closer to the point of origin.

Second, our investigation recommends a target width of 12 - 18 mm for blind users. In our study, target widths smaller than this range resulted in multiple overshootings (Figure 6), longer MT (Figure 4), and high error rate (Figure 5). This is different from existing guidance for sighted users, which suggests a minimum width of 7 to 10 mm [7, 24, 25, 31]. Naturally, this may affect the maximum number of items we can fit on a display. Designing layouts without considering this would lead to situations where some targets may not practically be selectable by blind users.

6.2 Application in accessible layout design

The new model can be used to identify optimal design parameters for layouts used in gliding-based selection. In this section, we demonstrate a case where the model guides a simple menu design for blind users. We enumerate design variables of interest, discretize them, and iterate over their possible values to find an optimum that minimizes the user's expected movement time for the menu.

Example: We take as an example a linear menu displaying 11 items from BlindSquare [10], a mobile application designed to aid blind users to navigate outdoors (Figure 12a). The mixture pointing model can predict the expected time for accessing an item given its height, based on a set of reasonable assumptions.

Assumptions. We assume that the user has a phone with a 104 mm tall display (e.g. iPhone 6), and navigates to the target from the top of the menu by gliding a finger on the screen. The menu can span over several screens, in which case we assume a 0.4 s cost for each screen change when navigating to later items in the menu.

Figure 12: An application example: Choosing design parameters to optimize selection time on a menu for gliding-based target selection. (a) Screenshot of BlindSquare application with 11 menu items. (b) and (c) Predicted average menu accessing time using mixture pointing model on this 11-item menu, under the assumptions declared in 6.2 for a frequent (b) and first-time (c) user. It shows that menu item heights of 13.9 mm and 12.1 mm lead to best performance for frequent and first-time users, respectively. The code for predicting the performance is provided in Supplementary Material.

Such value is an approximation of a 3-finger flickering (Apple’s VoiceOver [2]) or a 2-finger swiping (Android’s TalkBack [42]) for screen change on smartphones. If the user misses the target, a fixed penalty of 1.5 s will be introduced. This penalty was calibrated to the length of the duration of audio feedback in the experiment conducted in this study (see Figure 3). The assumed error rates are presented in Figure 5. Finally, we assume that the menu takes up 80% of the space on the display, as the screen also accommodates additional toolbars, as displayed in Figure 12a. We also distinguish a way of performing gliding for frequent and first-time users:

- **Frequent users.** We assume that a frequent user can approximate the location of the target and directly glide the finger to the item. For these users $MT_i = \Delta T_A + C + e \cdot P$, where ΔT_A is the movement time from the top of the menu to the top of the i -th item, at the distance A with item height H , which is modeled by the mixture pointing model (Equation 9) with parameters reported in Figure 11. The term C is the screen change cost in case the item is not located on the current screen, e is the probability of missing an item of height H and P is the penalty for wrong selections.
- **First-time users.** We assume that a first-time user would stop at each item to determine whether it is the intended target. Therefore, the time for accessing the i -th item on the menu $MT_i = (i - 1) \cdot (\Delta T_H + e \cdot P) + C$, where ΔT_H is the movement time traveling over a distance of a menu item height H , which can be modeled using the mixture pointing model. The term C is the screen change cost in case the item is not located on the current screen, e is the probability of

missing an item of height H and P is the penalty for wrong selections. We assume that a user can make selection errors at each item before reaching the intended target. Hence the error penalty weighted by probability ($e \cdot P$) is added for every item visited.

Predictions. Given the height (H) of a menu item, we plug in the mixture pointing model with parameter values in Figure 11 to predict the expected time of accessing menu items, assuming the distribution of each item being selected is uniform. Figure 12b shows the predicted accessing time for different item heights for a frequent user and Figure 12c for a first-time user. As shown, the optimized item height is 13.9 mm for frequent users, which would save approximately 0.13-0.41 s compared to using the item height suggestion of 7-10 mm for sighted users [7, 24, 25, 31]. For a first-time user, smaller item heights are favoured as they have to travel over the entire menu before reaching the target, with 12.1 mm being the optimal menu item height in this case. These findings support our recommended item size range for visually impaired users: item heights in range 12-18 mm, where the exact size depends on the display dimensions.

Software for menu optimization We provide a Python notebook for computing optimal item sizes for frequent and first-time users like in the example above. The notebook, attached to the paper in a supplementary file, can be used with different inputs, including screen height, screen change cost and penalty for errors.

6.3 Limitations and Future Work

Although the mixture model has substantially improved prediction accuracy over Fitts' law, increasing R^2 from 0.65 to 0.76, the prediction accuracy is still lower than for the pointing model for sighted users with visual feedback available. For example, previous work shows Fitts' law can robustly model pointing for sighted people with $R^2 > 0.9$ [33]. This indicates that blind users' pointing behavior is probably more complex than a mixture of two strategies. Our research is only the first step towards understanding this complex behavior.

In this paper, we focus on target selection tasks without distractors, which are similar to the basic Fitts-law-like pointing tasks. It is the first step towards understanding blind users' target selection behavior. In the future, it is worth investigating selection tasks with distractors. The idea of a distractor could be viewed as similar to the one defined in [9], except that the presence of the distractor would be auditory instead of visual. If distractors exist, a selection task may be viewed as concatenation of multiple basic gliding actions assuming the input finger will cross distractors before reaching the target (for novice users), or the user is aware of the approximate location of the target so that she can largely ignore the distractors (for expert users). To be more specific, we may estimate the target selection time for novice by aggregating the the estimated movement time from one distractor to another. For the expert, the gliding time estimation could be simply modeled as moving directly from the starting point toward the target. These assumptions are worth investigating in the future work.

We carried out the investigation using a particular design for the screen-reader, as explained in the Experiment section. The observed strategies may reflect this particular design, including the narration duration. Nevertheless, we implemented the screen-reader taking into consideration present services, such as the Android built-in Text-To-Speech service for narration. Many of the findings can be generalized beyond our particular settings.

7 CONCLUSION

We discovered that blind users adopt a mixed strategy when gliding the finger on a touchscreen to select targets. The mixture consists of two strategies: (1) iterative corrections to ballistic movements, relying on non-visual feedback, which is nonetheless well captured by Fitts' law, and (2) a previously unreported strategy where multiple sub-movements separated by stops are concatenated until the target is reached. Based on this finding, we proposed a mixture pointing model that relates the target selection movement time MT with movement amplitude A and target width W :

$$MT = a + bA + c \log_2\left(\frac{A}{W} + 1\right), \quad (16)$$

where a , b , and c are empirically determined parameters. This model reflects the mixture of the two strategies: The term $c \log_2(\frac{A}{W} + 1)$ captures the iterative-correction movement (Fitts' law), and the term bA reflects the multi-stop movement. A controlled experiment showed that this mixture model (Equation 16) outperforms Fitts' law in modeling MT : It increased R^2 from 0.65 (Fitts' law) to 0.76, and had higher prediction accuracy measured with information criteria which controlled for overfitting by adding penalty to the

extra parameter. It also outperformed a number of target selection models previously reported in the literature [13, 34, 35].

The mixture pointing model (Equation 16) advances our understanding of blind people's target selection performance: it reveals that the movement amplitude A plays a more important role than W in determining MT because of the multi-stop component bA , suggesting that shortening A is crucial for improving target selection performance of blind people. Our investigation leads to new implications for accessible user interface design, and the proposed model serves as a new tool to evaluate and optimize accessible user interfaces.

ACKNOWLEDGMENTS

We thank the anonymous reviewers for their insightful comments. This work was supported by NSF awards 1815514, 1805076, 1936027, NIH awards R01EY030085 and R01HD097188. AO and AP were funded by the Finnish Center for Artificial Intelligence FCAI, and the Academy of Finland projects Human Automata (Project ID: 328813) and BAD (Project ID: 318559).

REFERENCES

- [1] H. Akaike. 1998. Information theory and an extension of the maximum likelihood principle. In *Selected papers of hirotugu akaike*. Springer, 199–213.
- [2] Apple. 2020. Vision Accessibility in iPhone. <https://www.apple.com/accessibility/iphone/vision/>
- [3] S. Azenkot, J. O. Wobbrock, et al. 2012. Input finger detection for nonvisual touch screen text entry in Perkinput. In *Proceedings of graphics interface 2012*. 121–129.
- [4] P. Bach-y Rita, C. C. Collins, et al. Vision substitution by tactile image projection. *Nature* 221, 5184 (1969), 963–964.
- [5] G. Bailly, A. Oulasvirta, et al. 2014. Model of Visual Search and Selection Time in Linear Menus. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. Association for Computing Machinery, New York, NY, USA, 3865–3874. <https://doi.org/10.1145/2556288.2557093>
- [6] X. Bi, Y. Li, et al. 2013. FFitts Law: Modeling Finger Touch with Fitts' Law. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*. Association for Computing Machinery, New York, NY, USA, 1363–1372. <https://doi.org/10.1145/2470654.2466180>
- [7] X. Bi and S. Zhai. 2016. Predicting Finger-Touch Accuracy Based on the Dual Gaussian Distribution Model. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology (UIST '16)*. ACM, New York, NY, USA, 313–319. <https://doi.org/10.1145/2984511.2984546>
- [8] S. M. Billah, Y.-J. Ko, et al. 2019. Accessible gesture typing for non-visual text entry on smartphones. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–12.
- [9] R. Blanch and M. Ortega. 2011. Benchmarking pointing techniques with distractors: adding a density factor to Fitts' pointing paradigm. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 1629–1638.
- [10] BlindSquare. 2020. BlindSquare: Pioneering accessible navigation – indoors and outdoors. <https://www.blindsquare.com/>
- [11] K. P. Burnham and D. R. Anderson. A practical information-theoretic approach. *Model selection and multimodel inference* 2 (2002).
- [12] E. Crossman and P. Goodeve. Feedback control of hand-movement and Fitts' Law. *The Quarterly journal of experimental psychology. A, Human experimental psychology* 35 (06 1968/1983), 251–78. <https://doi.org/10.1080/14640748308402133>
- [13] A. D. Edwards. Modelling blind users' interactions with an auditory computer interface. *International journal of man-Machine studies* 30, 5 (1989), 575–589.
- [14] M. El Lahib, J. Tekli, et al. Evaluating Fitts' law on vibrating touch-screen to improve visual data accessibility for blind users. *International Journal of Human-Computer Studies* 112 (2018), 16–27.
- [15] P. M. Fitts. The information capacity of the human motor system in controlling the amplitude of movement. *Journal of experimental psychology* 47, 6 (1954), 381.
- [16] P. M. Fitts and J. R. Peterson. Information capacity of discrete motor responses. *Journal of experimental psychology* 67, 2 (1964), 103.
- [17] K. Gajos and D. S. Weld. 2004. SUPPLE: Automatically Generating User Interfaces. In *Proceedings of the 9th International Conference on Intelligent User Interfaces (IUI '04)*. Association for Computing Machinery, New York, NY, USA, 93–100. <https://doi.org/10.1145/964442.964461>

- [18] K. Z. Gajos, J. J. Long, et al. 2006. Automatically Generating Custom User Interfaces for Users with Physical Disabilities. In *Proceedings of the 8th International ACM SIGACCESS Conference on Computers and Accessibility (Assets '06)*. Association for Computing Machinery, New York, NY, USA, 243–244. <https://doi.org/10.1145/1168987.1169036>
- [19] K.-C. Gan and E. R. Hoffmann. Geometrical conditions for ballistic and visually controlled movements. *Ergonomics* 31, 5 (1988), 829–839.
- [20] K. Goldberg, S. Faridani, et al. A new derivation and dataset for Fitts' law of human motion. *EECS Dept., Univ. California, Berkeley, CA, USA, Tech. Rep. UCB/EECS-2013-171* (2013).
- [21] T. Guerreiro, K. Montague, et al. 2015. Blind people interacting with large touch surfaces: Strategies for one-handed and two-handed exploration. In *Proceedings of the 2015 International Conference on Interactive Tabletops & Surfaces*. 25–34.
- [22] L. He, R. Wang, et al. 2020. PneuFetch: Supporting Blind and Visually Impaired People to Fetch Nearby Objects via Light Haptic Cues. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems (CHI EA '20)*. Association for Computing Machinery, New York, NY, USA, 1–9. <https://doi.org/10.1145/3334480.3383095>
- [23] G. Hinton, I. Deng, et al. Deep Neural Networks for Acoustic Modeling in Speech Recognition: The Shared Views of Four Research Groups. *Signal Processing Magazine, IEEE* 29 (11 2012), 82–97. <https://doi.org/10.1109/MSP.2012.2205597>
- [24] C. Holz and P. Baudisch. 2010. The Generalized Perceived Input Point Model and How to Double Touch Accuracy by Extracting Fingerprints. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*. ACM, New York, NY, USA, 581–590. <https://doi.org/10.1145/1753326.1753413>
- [25] C. Holz and P. Baudisch. 2011. Understanding Touch. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. ACM, New York, NY, USA, 2501–2510. <https://doi.org/10.1145/1978942.1979308>
- [26] J. Hong, A. Pradhan, et al. 2017. Evaluating Wrist-Based Haptic Feedback for Non-Visual Target Finding and Path Tracing on a 2D Surface. In *Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '17)*. Association for Computing Machinery, New York, NY, USA, 210–219. <https://doi.org/10.1145/3132525.3132538>
- [27] S. K. Kane, J. P. Bigham, et al. Fully accessible touch screens for the blind and visually impaired. *University of Washington* (2011).
- [28] S. K. Kane, J. P. Bigham, et al. 2008. Slide rule: making mobile touch screens accessible to blind people using multi-touch interaction techniques. In *Proceedings of the 10th international ACM SIGACCESS conference on Computers and accessibility*. ACM, 73–80.
- [29] S. K. Kane, M. R. Morris, et al. 2011. Access Overlays: Improving Non-Visual Access to Large Touch Screens for Blind Users. In *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology (UIST '11)*. Association for Computing Machinery, New York, NY, USA, 273–282. <https://doi.org/10.1145/2047196.2047232>
- [30] S. Keele. Movement Control in Skilled Motor Performance. *Psychological Bulletin* 70 (12 1968). <https://doi.org/10.1037/h0026739>
- [31] G. LLC. 2020. Accessibility in Material Design. <https://material.io/design/usability/accessibility.html>.
- [32] I. S. MacKenzie. A note on the information-theoretic basis for Fitts' law. *Journal of motor behavior* 21, 3 (1989), 323–330.
- [33] I. S. MacKenzie. Fitts' law as a research and design tool in human-computer interaction. *Human-computer interaction* 7, 1 (1992), 91–139.
- [34] D. E. Meyer, R. A. Abrams, et al. Optimality in human motor performance: ideal control of rapid aimed movements. *Psychological review* 95, 3 (1988), 340.
- [35] D. E. Meyer, J. E. Keith-Smith, et al. 1990. Speed-accuracy tradeoffs in aimed movements: Toward a theory of rapid voluntary action.
- [36] J. Müller, A. Oulasvirta, et al. Control theoretic models of pointing. *ACM Transactions on Computer-Human Interaction (TOCHI)* 24, 4 (2017), 1–36.
- [37] A. Oulasvirta, N. R. Dayama, et al. Combinatorial optimization of graphical user interface designs. *Proc. IEEE* 108, 3 (2020), 434–464.
- [38] M. Ptito, S. M. Moesgaard, et al. Cross-modal plasticity revealed by electrotactile stimulation of the tongue in the congenitally blind. *Brain* 128, 3 (2005), 606–614.
- [39] E. Ricciardi, D. Bonino, et al. Do we really need vision? How blind people “see” the actions of others. *Journal of Neuroscience* 29, 31 (2009), 9719–9724.
- [40] S. Sarcar, J. P. Jokinen, et al. Ability-based optimization of touchscreen interactions. *IEEE Pervasive Computing* 17, 1 (2018), 15–26.
- [41] A. Savitzky and M. J. Golay. Smoothing and differentiation of data by simplified least squares procedures. *Analytical chemistry* 36, 8 (1964), 1627–1639.
- [42] TalkBack. 2020. TalkBack: An Open Source Screenreader For Android. <https://support.google.com/accessibility/android/answer/6283677>
- [43] S. Watanabe. Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. *Journal of Machine Learning Research* 11, Dec (2010), 3571–3594.
- [44] A. T. Welford. The measurement of sensory-motor performance: Survey and reappraisal of twelve years' progress. *Ergonomics* 3, 3 (1960), 189–230.
- [45] A. T. Welford. 1968. *Fundamentals of skill*. Methuen, New York, NY, USA.