One-thumb Text Acquisition on Force-assisted Miniature Interfaces for Mobile Headsets

Lik-Hang LEE*, Yiming ZHU†, Yui-Pan YAU†, Tristan BRAUD†, Xiang SU‡* Pan Hui†‡

*Center for Ubiquitous Computing, University of Oulu, Oulu, Finland
†Department of Computer Science and Engineering, Hong Kong University of Science and Technology, Hong Kong
‡Department of Computer Science, University of Helsinki, Helsinki, Finland
lhleeac@connect.ust.hk, braudt@ust.hk, arthur.yau@connect.ust.hk, panhui@ust.hk

Abstract—**Touchscreen interfaces are shrinking and even disappearing on mobile headsets. The existing approaches for text acquisition on mobile headsets, for instance, speech commands and hand gestures, are cumbersome and coarse. In this paper, we show the feasibility of interaction on a miniature area as small as 12 * 13 mm^2 that offers an input alternative on small form-factor devices such as smartwatches, smart rings, or the spectacles frames of mobile headsets. To this end, we propose and implement two interaction approaches, namely FRS and DupleFR, for acquiring textual contents on mobile headsets. Both approaches leverage force-assisted interaction on a miniature-size interface. They enable the user to acquire textual content with various granularities such as characters, words, sentences, paragraphs, and the entire text. After 8 sessions, 22 participants with FRS and DupleFR achieve the peak performance of respectively 11.455 and 10.611 seconds per textual acquisition with accuracy rates of 91.41% and 94.95%. Although FRS and DupleFR as indirect manipulations are disadvantageous, they are at least 37.06% faster than the commercial standards designated to direct manipulation on touchscreens.**

I. INTRODUCTION

Mobile headsets enable users to interact with diverse content, in the form of text, images, videos, and 3D objects appearing in an integrated physical and digital environment. Many academic works study interaction with digital overlays of images and videos on the top of the physical world. However, most studies neglect interaction with text content through mobile headsets. On the other hand, the existing commercial approaches for interaction with text content are indirect and ineffective. Users take photos to acquire text information, for example, shown on a poster or advertisement[1]. Furthermore, there exists no practical way for users to perform direct manipulation on text information in mobile headsets' web browsers. It is thus necessary to develop systems and techniques for text interaction on mobile headsets.

On mobile devices featuring a touchscreen, tap, hold and drag gestures allow users to select the text content directly. Due to the limited size of mobile headsets, touchscreens are no longer available, and users are unable to touch text elements such as characters and words directly. Even with voice assistants (e.g. Cortana on Microsoft Hololens), difficulties arise because of the awkward positioning of the targeted text elements. In a thread of the Microsoft Developer Forum for Microsoft Hololens[2], users reflect the need for a Bluetooth-paired keyboard and mouse to manipulate text content. These bulky external controllers are designed for sedentary desktop scenarios and thus deteriorate the mobility of mobile headsets. In addition, mid-air interaction with the depth camera on Microsoft Hololens is coarse and prevents accurate targeting. Prolonged mid-air interaction is subject to the Gorilla Arm Syndrome, where the users require lifting their arms. Considering both the mobility and usability, a small-size interface for off-hand posture should be designed for conveniently acquiring textual contents.

This paper first explores three reachability techniques for acquiring characters in dense and cluttered textual contents. The three techniques include one commercial standard (Keyboard Trackpad) and two state-of-the-art academic solutions (BezelCursor [16] and ForceRay [15]). Among the techniques, ForceRay, a force-assisted ray casting technique, demonstrates a prominent advantage in terms of space-saving in the interaction area, which is at least 8.3 times smaller than other alternatives. By leveraging force-assisted interaction, the small-size interaction space can meet the emerging demand for small-size smartglasses designed for mobile scenarios. Accordingly, this paper presents two multi-step approaches driven by force-assisted interactions, named ForceRaynSelect (FRS) and DupleForceRay (DupleFR), for text acquisition at various granularities such as character, word, sentence, paragraph, as well as the entire text.

In order to preserve the small-size interface, force-assisted interaction serves as the backbone for these multi-step interaction methods as follows. In FRS, the user first exerts force on the thumb-sized button to position the cursor at the target character. The cursor moving distance is directly proportional to the applied force level. Next, the user chooses a text acquisition unit by varying the force level, ranging from character level to the entire document. Finally, the user can select text through circular gestures (rotations) on the touch interface. In DupleFR, the user exerts force on the thumb-sized button to position the cursor at the first target character as the head of the selection. The user then repeats the first step to target another character as the tail of the selection. Finally, the user confirms the selection between the head and tail

---

characters. We implement the proposed interaction approaches (FRS and DupleFR) on an iPhone 7. The circular button interface features a diameter of 10.5 mm centred at the (x,y) coordinate of (187.5, 475.5) on the smartphone touchscreen (bottom-centre). The user can acquire textual contents with this button through both multi-step approaches on the distal screen of smartglasses. The existing implementation enables the user to accomplish the tasks with one-thumb interaction with off-hand posture in a concealed manner.

Across all experimental trials, a total of 22 participants were able to complete a text acquisition task with FRS and DupleFR in 14.009 and 11.633 seconds with average accuracy rates of 87.69% and 94.26%, respectively. During the final trial, the participants achieved average completion times of 11.455 and 10.611 seconds with the peak accuracy rates of 91.41% and 94.95%. Even though FRS and DupleFR pose the disadvantage of indirect manipulation, they are 37.06% and 41.70% faster than the commercial standards of direct manipulation on touchscreens. Among the two approaches, the majority of participants show a significantly higher preference for DupleFR than FRS. A miniature area of 153 pt * 170 pt on the iPhone 7 screen (326 ppi), corresponding to 12 * 13 mm^2 in real life can accommodate the proposed approaches. Small-size gadgets such as smart rings and smart jewellery can thus apply FRS and DupleFR on thumb-sized interfaces for text acquisition. To the best of our knowledge, this paper presents the first text acquisition system for mobile headsets, which reduces the burden of selecting text contents on mobile AR scenarios. We propose and investigate two force-assisted approaches. FRS and DupleFR introduce a subtle interaction approach on minimalist interfaces, yielding the following contributions and improvements over other existing systems:

1) This paper presents the first solution to improve the performance of text acquisition at various granularities on mobile headsets.
2) The investigation of reachability techniques show the potential of force-assisted interaction for text acquisition.
3) FRS and DupleFR define a miniature interaction area (12 * 13 mm^2) for text acquisition than other state-of-the-art solutions, which satisfy the needs of mobility and usability.
4) The proposed solutions enable off-hand interaction without noticeable action, and allows for extending the input bandwidth (input capability) of smart wearables.

The rest of this paper is organized as follows. We first summarize the major related studies in Section II. We then conduct a pilot test to explore the space requirement with different interaction techniques in Section III, and introduce our FRS and DupleFR, our interaction approaches in Section IV. Finally, we describe our implementation and evaluate the proposed solution in Section V.

II. RELATED WORK

The proposed force-assisted interaction approaches are related to several topics, among which one-handed gestures,

object manipulation on smartglasses as well as leveraging force as a modality.

A. One-handed Target Acquisition with Unreachable Targets

In this paper, we consider Target Acquisition as the action of selecting icons and menus in an interface, character keys on a keyboard, or 3D objects in Virtual and Augmented Reality interfaces on mobile headsets. The research community in Human-Computer Interaction (HCI) has proposed various gestures and interfaces to acquire targets indirectly.Smartphones with screens beyond 5" limit the reach of the user's thumb with one-handed operations. For one-handed operations to remain possible, most existing smartphones use a transformable screen interface, for instance, Samsung's smartphones scale the screen down to two-third of its original size after triple-tapping the home button. The research community has proposed various gestural techniques to trigger the screen transformation to enable users to shift the interfaces toward the thumb reach. Such techniques include tilting the smartphones [17], sliding the screen [18], swiping at the edge area of the screen [19], touching the additional sensory hardware at the rear of smartphones [20], and sideways keyboard [21]. Although these techniques move the target closer to the thumb, they still prevent access to the areas that are physically outside of the user's reach. Other methods suggest extending the thumb reach through gestures that change the size [22] or direction [16] of the pointing device.

Smartglasses share similar problems with large-size touchscreen mobile devices. In this setting, the embedded small-size interfaces (trackballs, mini-touchpads, buttons) do not provide full functionality compared to desktop computers [23]. The users cannot directly reach the digital contents shown on the smartglasses displays [25]. Although speech recognition is one of the significant interaction methods on smartglasses, it might be vulnerable in a shared or noisy environment [27]. Furthermore, speech commands are less preferred than body gestures and handheld devices input [28]. Our work aims to design gestures on a minimalist interface to be embedded into a smart ring for enhancing user mobility [5]. Even though smartphone screens have been the object of numerous studies, only a few works consider untouchable displays such as smartglasses. In this paper, we investigate one-handed gestures on minimum-size interfaces to reach targets on untouchable displays.

B. Objects Manipulation on Smartglasses

Various hand gesture-based interfaces exist for interaction with mobile headsets [23]. These interfaces target various goals, for instance, facilitating digital contents in office environment [29], moving 3D objects [11], facilitating interactions with digital contents in office environment [29], mincing fiducial marker by hand gestures [30], text input with mid-air keyboards [10], as well as 6DOF movements of 3D objects [14]. However, the research on text acquisition on smartglasses remains limited. Hand gesture-based interfaces are subject to coarse detection [24], which are not appropriate

for selecting small and dense targets such as textual contents. In contrast, touch interfaces [7] and additional sensors [12] enable responsive and accurate target acquisition, such as selecting the textual contents down to character level.

Text acquisition is a commonplace on touchscreen interfaces, and prior researches have proposed many techniques to manipulate textual contents for copy-and-paste operations, for instance, two-handed touch and drag [4], direct tap with word-level snapping [2], tap-to-edge and slide-to-edge gestures to capture text across pages [6]. With the goal of redesigning touch-based gestures for indirect text acquisition on smartglasses, our paper is the first effort to address text acquisition on smartglasses where direct touch gestures are inapplicable. Under this constraint, our proposed system enables users to select textual contents in different granularities, including characters, words, paragraphs, and entire documents, on smartglasses.

C. Force Interactions on mobile devices

Humans can distinguish between subtle amounts of force ranged from 6 [33] to 10 [34] levels. Force as a modality can distinguish choices through varying the strength of force exertion in a continuous spectrum. ForceEdge [3] enables scrolling a list of items (e.g. date) through exerting force on a force-sensitive touchscreen instead of performing a swipe gesture on the scrolling items. Goguey et al. [1] apply force-augmented taps on a touchscreen to alter the granularity of text content selection. A recent study proposes a force-driven cursor allowing users to get the items located outside of thumb reach [15]. Force as a threshold can trigger a mode switching in the interfaces, for example, applying force to input uppercase letters without tapping the 'shift' key [31]. Force can also be applied for distinguishing multiple characters in an ambiguous keyboard, such as two characters [32] or three characters [7] in one key. However, none of these methods leverages the advantages of force input to tackle text acquisition with one-handed gestures on smartglasses. To the best of our knowledge, we are the first work to explore the force modality and construct a thumb-size interface to acquire the non-touchable textual contents on mobile headsets.

III. ONE-THUMB REACH ON TEXTUAL CONTENTS

In this section, we investigate three candidate gestures for one-handed target acquisition. These candidates include both a commercial standard and two state-of-the-art academic solutions. The three selected gestures are Keyboard Trackpad Mode (KT)[3], BezelCursor (BC) [16] and ForceRay (FR) [15]. KT is a common technique on commercial platforms and serves as our baseline. BC is one of the first mobile techniques with solely touch modality, and is frequently used as a benchmark in other studies. FR is the most recent technique leveraging force as the key modality. However, the study only focuses on reaching an array of large icons (mixed with

---

[3]"How to use keyboard trackpad model on every iPhone and iPad with iOS 12" https://9to5mac.com/2018/09/17/how-to-use-keyboard-trackpad-mode-on-every-iphone-and-ipad-with-ios-12/
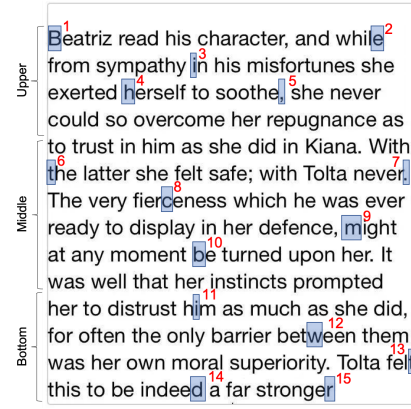

Fig. 1. Testing interface of the first user study (target character 1 – 15)

distracting icons) out of thumb reach and neglects more closely compacted targets such as textual content.

A. Participants and Apparatus

To understand users' behaviour with the candidate gestures, we conduct a user study with 13 participants (19 - 29 years; 8 male and 5 female; all right-handed), recruited on the university campus. All of them are experienced smartphone users. Only one of them has prior experience with force interaction.

We implement the three interaction techniques on an iPhone 7 equipped with a pressure-sensitive touchscreen (4.7"). In addition to the touch modality, the touchscreen is responsive to pressures ranging continuously from 0 to 3.3 N, corresponding to 0 to 6.66 units in iOS (Swift) programming [35]. The touch and force modalities of the iPhone 7 can entirely support the three gestures.

Figure 1 shows the testing interface. We dedicate the bottom part of the screen to the pointing method implementation. The rest of the screen presents a text extracted from a novel named *Kiana: a Tradition of Hawaii - Chapter XIV* written by Jarves et. al [4]. We display the text using the *Helvetica Neue* font in a font size of 18-pt [41]. All interaction methods allow moving a cursor over the text for character selection.

B. Techniques

In this pilot experiment, we target three state-of-the-art techniques for character selection. We represent these techniques on Figure 2.

a) Keyboard Trackpad (KT): This technique mimics the iPhone's trackpad mode on the iOS keyboard. The user can switch between the QWERTY keyboard and a trackpad by deep-pressing anywhere on the keyboard. The user then keeps the force level higher than the threshold value (1.33 out of 6.66) to maintain the Trackpad mode, and controls the position of the caret (text cursor) on the screen through pan gestures. The touch coordinates of the pan gesture map with the caret through geometric transformations from the trackpad area

---

[4]Online version of Kiana: a "Tradition of Hawaii" https://quod.lib.umich.edu/p/philamer/AJA1762.0001.001?view=toc
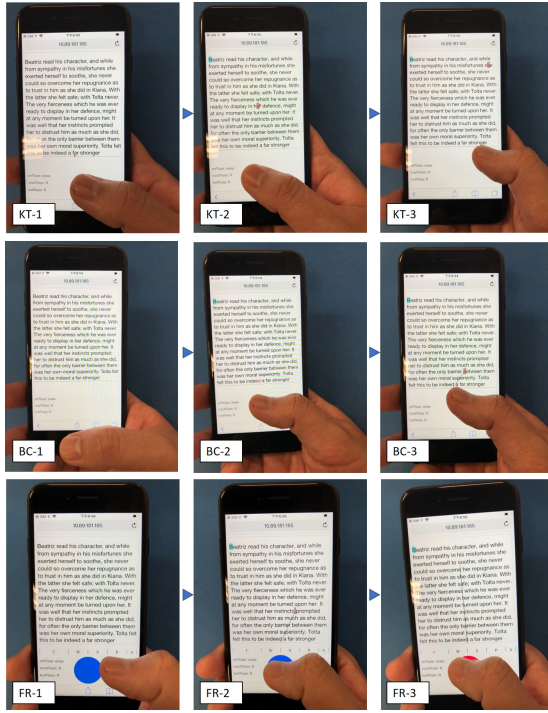
Fig. 2. The three interaction techniques (Top – Keyboard Trackpad (KT), Middle – BezelCursor (BC) and Bottom – ForceRay (FR)): KT-1: the thumb does not touch the screen, the cursor is at an idle location (bottom edge); KT-2: the thumb presses over the threshold. The cursor appears and maps with the thumb touch coordinate; KT-3: the thumb moves to the upper right corner. The cursor follows; BC-1: the thumb does not touch the screen and no BC appears; BC-2: the thumb taps on the touchscreen and activates a cursor; BC-3: a swipe gesture moves the cursor by approximately 45°; FR-1: the thumb does not touch the screen. No cursor and no FR appear; FR-2: the thumb presses over the threshold and a cursor towards 45° appears; FR-3: a harder press moves the cursor to the farther edge. The thumb moves to adjust the cursor angle to approximately 0°.

(landscape rectangle of 356 * 160 pt) to the caret interaction area (portrait rectangle). The user can move his or her thumb in all directions to reach the character target. When the caret reaches a target character, the character's background changes colour. Lifting off the thumb confirms the character selection.

b) BezelCursor (BC): This technique casts a pointing line from the centre of the interaction area. A rectangle cursor at the end of the line represents the extended thumb reach on the screen. To initiate the BezelCursor, the user swipes from a bezel to the central area of the touchscreen, where a landscape rectangle of 356* 160 pt at the bottom bezel serves as the reference starting point. The swipe gesture triggers a pointing line that extends proportionally to the thumb displacement and rotates based on the thumb direction. The line grows proportionally to the thumb displacement using a 1:3 ratio [16], a comfortable ratio for most users. The original implementation uses the DynaSpot cursor [36] that expands depending on the swiping acceleration. The cursor also seeks the closest icon. However, this approximation is inappropriate for dense and equally distanced textual contents. We thus disable the DynaSpot cursor to avoid selecting neighbouring targets. Similarly to the Keyboard Trackpad, the intersection
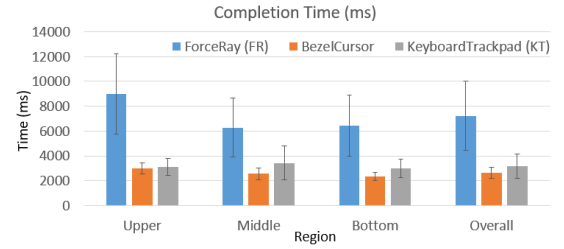

Fig. 3. Completion Time (in ms) for three gesture candidates for one-handed target acquisition; A higher value means the technique is more time consuming.

of the cursor with the text leads to highlighting a character and lifting the thumb selects the target character.

c) ForceRay (FR): This technique leverages the 3D Force touch on iPhone touchscreens to achieve a series of subtle interaction procedures. The user activates FR by pressing on the touchscreen with a force level over a threshold of 1.33 units [15]. To keep FR activated, the user needs to maintain a force above the threshold on a button of diameter 10.5 mm centred at x,y = (187.5, 475.5). Similarly to the BezelCursor, the system casts a ray ended with a rectangle cursor as an extension of the user's thumb. The user's force level can control the ray's length – the higher the force, the farther the ray extends. A reduced force accordingly leads to the cursor coming back to the thumb side. The ray's length grows linearly with the applied force, ranging between the threshold value and the maximum force level [38]. Besides, the thumb rotation, computed by the displacement from the first touch coordinate (reference point), steers the ray orientation. The intersection of the cursor with the text highlights a character. As the force level gradually changes when lifting the thumb to select the target character, we apply the Quick Release technique [37] for character selection. The system looks up the characters selected between 200 and 300 ms before lifting the thumb off the screen and selects the most frequent character.

C. Task

We ask the participants to complete the same task for each of the three candidate methods. A task consists of using the designated method to target 15 characters in the text. We divide the screen into three regions: upper, middle and bottom. Each region contains five character targets. The 15 target positions are widespread on the screen in order to evaluate the limits of each technique's reachability. During each trial, the targets are be highlighted one by one following a sequence from 1 to 15. As we aim to understand the user behaviour in this pilot test, each participant only completes one trial (15 targets) with each technique. We tell the users to select the target character as fast and accurately as possible.

D. Results

1) Performance: Figure 3 shows the completion times for the three techniques. The error bars represent the standard deviation. Regarding the completion time, two-way repeated measures ANOVA yields a significant effect of the interaction
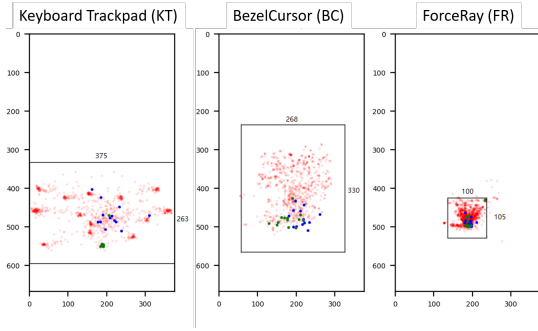
Fig. 4. Touch footprint for the three techniques (captured every 20 ms): red dot as x and y coordinate on the touchscreen; blue dot as the starting point of the technique; green dot as the ending point of the technique; black box as the interaction area.
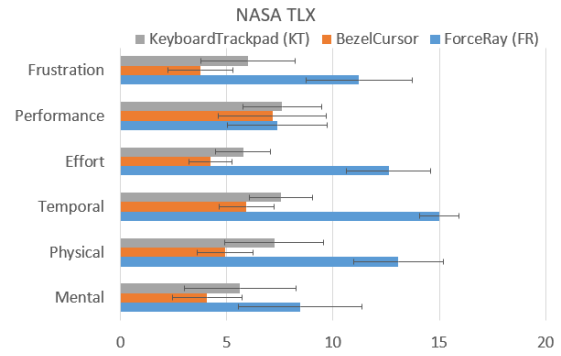

Fig. 5. NASA Task Load Index (TLX) survey results for the three candidate gestures for one-handed target acquisition; A higher value means the technique is more demanding.

technique and the locations of target characters ($F_{2,14}$ = 131.9015, p <0.0001), indicating that different one-thumb reachability techniques result in different performances, accompanied by an effect of different locations on the testing interface. Participants achieve an average completion time of 7.236 s (σ= 5.562 s), 2.628 s (σ= 0.882 s), and 3.181 s (σ= 1.932 s) with FR, BC, and KT respectively. The users report that selecting characters with FR in the upper region is less reliable than other regions. The average completion time in the upper region (8.984 s (σ= 6.474 s)) is at least 39% higher than the middle (6.271 s (σ= 4.771 s)) and bottom (6.452 s (σ= 4.934 s)) regions. Users are indeed susceptible to overshoot due to the limitation of the thumb's first carpometacarpal joint. Due to this effect, users are less capable of controlling the FR when applying maximum force on the screen [38].

Regarding the accuracy, two-way repeated measures ANOVA barely shows a statistical significance of the interaction techniques and the locations of target characters ($F_{2,13}$ = 3.0625, p= 0.0475). Among the 630 targets (210 characters * 3 techniques), FR, BC and KT respectively achieve an accuracy of 92.86%, 96.67% and 90.95%. Surprisingly, FR outperforms the pan-based KT and performs slightly worse than swipe-based BC.

2) Interaction Footprint: Figure 4 shows the gesture footprint generated by each technique on an iPhone 7 with a 375 pt * 667 pt screen. To improve the readability, the red, green and blue dots shown in Figure 4 are footprint captured every 500 ms, while we compute the interaction areas (black box) using samples captured every 20 ms. We calculate the interaction area by first finding the centre of the captured samples. We then compute the standard deviation on each axis. The interaction area corresponds to three times the standard deviation from the centre (99.7% of the samples). If the calculated interaction area exceeds the screen boundaries, we consider the screen edges as the interaction area boundaries for the considered edge. KT, BC and FR respectively produce interaction areas of 375 pt * 263 pt, 268 pt * 330 pt and 100 pt * 105 pt. Among the three techniques, FR produces the smallest interaction area and the densest footprint, which is respectively 9.4 times and 8.3 times smaller than KT and BC. This smaller area allows the participants to keep their

interactions within the most comfortable thumb reach with a firm grip.

3) NASA TLX: After the pilot tests for the three techniques, we ask the users to complete a NASA Task Load Index (NASA TLX [39]) form for evaluating their perception of the load induced by the experiment for each candidate gesture. This survey requires the user to grade six qualitative metric ranging from 1 to 20. These metrics include Mental Demand, Physical Demand, Time Demand, Effort, Performance, and Frustration. We display the results of the survey in Figure 5. One-way ANOVA under three techniques shows an effect of the technique for all metric (p <0.05) except for the Mental Demand (p-value = 0.0833) and the Performance (p-value = 0.09668). These results imply that the participants perceive no difference between the three gesture candidates in terms of Mental Demand and Performance. The Bonferroni and Holm methods show that the participants perceived no difference in Physical demand (Bonferroni p-value = 0.43), Time (Bonferroni p-value = 0.32), and Frustration (Bonferroni p-value = 0.58) between the Keyboard Trackpad (KT) and the BezelCursor (BC). On the other hand, FR displays 50% – 120% higher load scores than KT and BC for all metric except the performance metric. The participants agree that the four above aspects of ForceRay (FR) are significantly different from KT and BC. From the user rating, we can conclude that the participants perceived a significantly higher load using FR than KT and BC. The key reason is that the majority of them had no experience with force-assisted interaction, and FR needs multiple training sessions to become skilful (up to 12 training sessions [15]). On the other hand, the participants are very familiar with pan (KT) and swipe (BC) gestures.

4) Design Implications: In this paper, our primary goal is to design a minimalist interface. The prominent feature of FR is the small-size interaction area, which is at least 8.3 times smaller than BC and KT. The completion time and user perception of FR are less advantageous than BR and KT, as the participants are very skilful at pan and swipe gestures (KT and BC) but have no experience to drive the cursor under a continuous force spectrum. However, FR has the potential to become a viable option, as more testing trials can compensate

for the performance gap and hence the user perception [15]. The force interaction serves as an alternative to the spacious interaction for swipe and pan gestures, which shed light to designing novel interaction approaches for small form-factor wearable computers. In the next section, we design two interaction techniques on the basis of force-assisted interaction.

IV. DESIGN OF INTERACTION APPROACHES

In this section, we discuss the design of two interaction approaches. These approaches, respectively named ForceRaynSelect (FRS) and Duple ForceRay (DupleFR), are based on the chosen ForceRay (FR) technique. Both approaches share the following five features: 1) their interaction areas inherit the characteristic of FR and can fit within a button sized at 10.5 diameters and centred at (x,y))=(187.5, 475.5) on a 375 pt * 667 pt iPhone screen; 2) both techniques can acquire the textual contents with full coverage of granularity, from a character to the entire text; 3) both of them leverage force-assisted interaction to enable multi-steps interaction;4) the multiple steps in each approach respond to the needs of reachability and textual granularity to select text when directly touching the content is no longer possible; 5) one-thumb selection happens and Quick Release [8] support the force-assisted approaches.

A. Approach 1: FRS

Figure 6 depicts FRS, our first interaction approach, consisting of three fundamental steps.

a) Step 1 – Targeting Character: First, the user applies ForceRay to control the cursor and reach a target character.

b) Step 2 – Selecting granularity: The user then selects the textual granularity in the continuous force spectrum with ForceSelect (FS). Granularity levels include character, word, sentence, paragraph as well as entire text. These granularities are displayed within a horizontal slider above the main control area. Under the force spectrum from 0 to 6.66 unit, the ratio of character (C), word (W), sentence (S), paragraph (P) and the entire text (A) are respectively 4:2:2:2:1.

c) Step 3 – Text selection: Finally, after selecting the granularity, the user performs a non-force sensitive circular gesture on a bi-directional button. Each 45° selects an additional textual unit. Taking the selected character as a reference point, the user draws circles either clockwise or counter-clockwise to select the textual content positioned after or before the reference point with the given granularity. The extent of text selection depends on the distance of circular gestures. The button represents 8 equal size division, and a circle is equivalent to 8 textual units in the chosen granularity. As such, one circle in word level selects 8 word, while 1.5 circles in character level represent 12 characters.

The user confirms the decision for each step by performing a tap on the button. During each step, in case of an unwanted decision, the users can do a 'Hard Press' to revert the decision (character target/ granularity/ number of textual contents). The 'Hard Press' is triggered when the maximum force is reached within 200 ms after an incorrect decision.
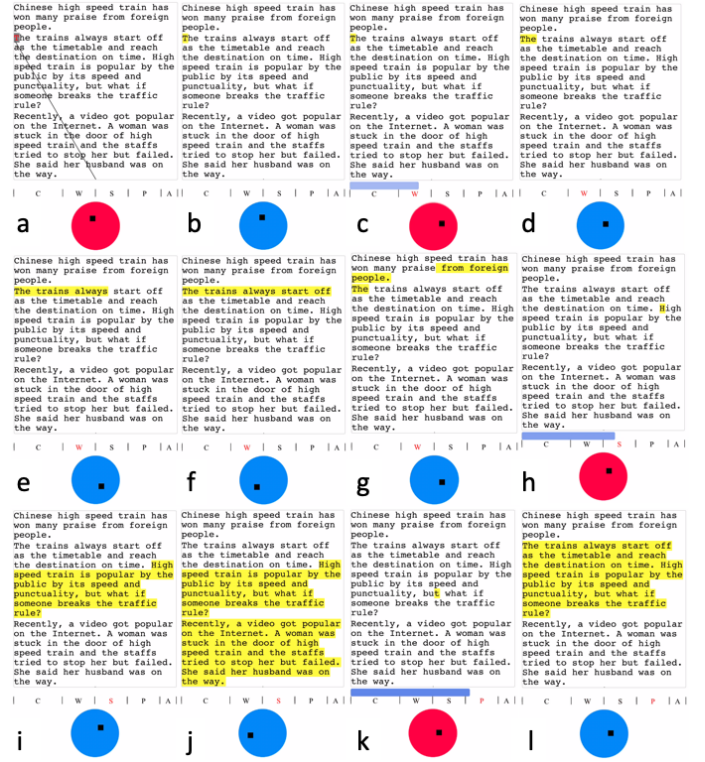

Fig. 6. An illustration of FRS. The Black dot represents the user's touch position. The Button changes color depending on the whether the force level exceeded a threshold level (red) or not (blue). **(a)** Apply FR to reach a character target (step 1); **(b)** A character is chosen, a tap confirms the character selection (step 1); **(c)** Apply FS to set the selection granularity (step 2); **(d)** Word level is chosen, a word is selected (step 2); **(e)** Circular gesture in clockwise direction selects more words to the right (step 3); **(f)** Further circular gesture in clockwise direction (step 3); **(g)** Circular gesture in counter-clockwise direction, inverts the selection (step 3); **(h)** Apply FS to select the textual granularity (step 2); **(i)** Sentence level is chosen, a sentence is selected (step 2); **(j)** Circular gesture in clockwise direction to choose additional sentences (step 3); **(k)** Apply FS to select the textual granularity (step 2); **(l)** Paragraph level is chosen and a paragraph is selected (step 2).

We justify our design choices as follows. The first step selects the target through the FR technique as the touchscreen is not available. The second step is modified from the work of ForceSelect (FS) [1], which enables users to select the granularity of textual contents through performing a pan gesture with a force level specifically for the targeting granularity on the touchscreen. However, mobile headsets do not feature touchscreens, and thus the third step of circular gestures substitutes the pan gestures.

B. Approach 2: DupleFR

In comparison to FRS, DupleForceRay (DupleFR) only relies on two steps, as shown in Figure 7. DupleFR leverages the granularity of textual contents where the fundamental element is character level. Instead of fixing the granularity as in FRS, this technique picks the head and tail of the textual body. The user performs two successive ForceRay to reach the head and tail target characters. Similarly, the two sequential steps will proceed when the user performs a tap on the button
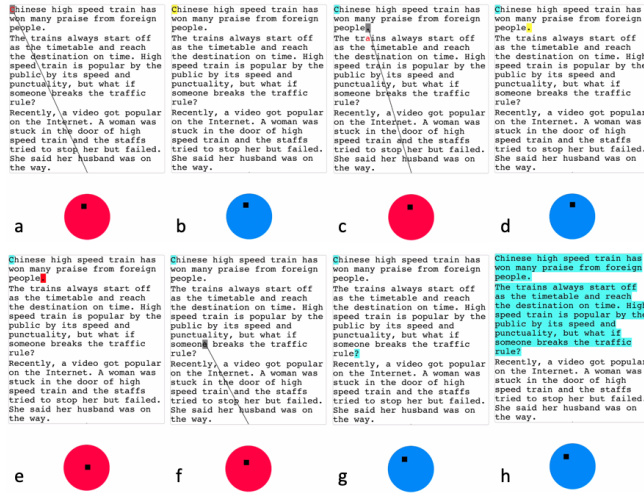
Fig. 7. An illustration of DupleFR, where Black dot is the user's touch position, Buttons of either in red color or in blue color indicate the force level exceeded threshold level or not. (a) Apply FR to reach a character target at the head position (step 1); (b) A character is chosen, a tap confirms the character selection (step 1); (c) Apply FR to reach another character target at the tail position (step 2); (d) A character is selected (step 2); (e) Hard Press to revert the chosen character (step 2); (f) Apply FR to choose a new character target at the tail position (step 2); (g) A new character is chosen, a tap confirms the final selection (step 2); (h) The selected textual contents.
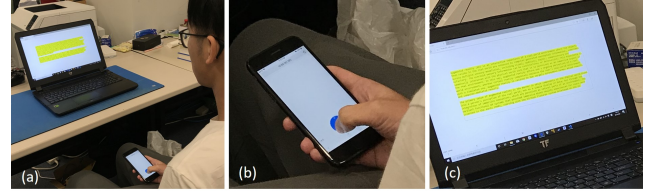

Fig. 8. (a) The evaluation setup; (b) a force-assisted button interface on an iPhone 7 to select the textual contents shown on (c) a distal display.
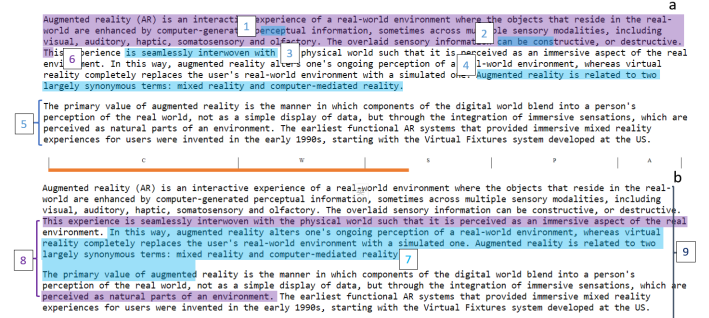

Fig. 9. An example of 9 textual types in the user evaluation interface: (a) Full testing interface at 16:9 ratio where a force spectrum for FRS located at the edge area; (b) an illustration of other tasks in the same interface.

to confirm the decision. In case of an unwanted decision, the users can do a 'Hard Press' to cancel the selected character target in each step.

V. USER EVALUATION AND DISCUSSION

In this section, we evaluate the user performance of FRS and DupleFR. We recruited 22 participants (19 - 33 years ($\overline{M} = 25.82, \sigma=3.94$); 21 male/1 female; 21 right-handed) from the university campus. All of them are experienced smartphone users but have no prior experience with force interaction. Each participant attended an approximately 70-minute experiment. For each technique, the participants ran over 8 sessions, and a total of 461,720 and 263,840 active touchpoints for FRS and DupleFR were recorded. After a 10-minute warm-up session, the participants are instructed at the beginning of each session to complete the task as fast as possible. They can only correct their mistakes for the current sub-tasks.

A. Apparatus and Tasks

The user evaluation aims at understanding the user behaviour and user performance of the two proposed interaction approaches under various textual granularity [1]. Participants (Figure 8) in a seating posture hold an iPhone 7 to acquire the textual contents shown on a distal 16:9 screen (15") of a laptop emulating the Microsoft Hololens display. As some participants cannot read the contents clearly on the mobile headset [40], a computer screen was used instead, which guarantees the users to read the dense and small textual contents clearly. The visual clues from ForceRay and ForceSelect are shown on the laptop, while only a thumb-size circular button (diameter = 10.5 mm, centred at (x,y) coordinate = (187.5, 475.5)) remains on the touchscreen of the smartphone.

As shown in Figure 9, every session consists of 9 sub-tasks considering various textual granularities, as follows: (1) Sub-word as 4 characters inside a word; (2) Word and character as 2 words and 4 characters; (3) 4 words; (4) 1 sentence; (5) 1 paragraph; (6) 2 sentences and 2 characters; (7) 2 sentences and (across paragraph) 5 words; (8) 4 sentences (across paragraph); (9) Entire text. The material is extracted from an online publishing platform named Medium[5], which is regarded as a professional and easy-to-understand writing style, so the tasks are reasonable mock-ups of everyday reading material. We apply a font size of 18-pt [41] in all sessions.

B. Results

1) Completion Time: Two-way repeated measures ANOVA demonstrates a significant effect of the Interaction Approaches and the Sessions ($F_{1,7} = 57.06$, p <0.05), which indicates the significance of interaction approaches on the completion time and the learning effect between sessions on similar interaction approach conditions. Figure 10 (left) shows the average completion times for FRS and DupleFR over the course of the 8 sessions. On average, over the 8 sessions, the participants with DupleFR (avg=104.697 s, $\sigma=23.607$ s) complete the tasks 21% faster than FRS (avg=126.081 s, $\sigma=35.217$ s). During the first session, we notice a large difference between both methods. The complexity of FRS compared to DupleFR leads to participants taking more time to complete the task, 167.434 s ($\sigma=39.830$) compared to 117.231 s ($\sigma=26.454$). At the eighth session, the participants improve their completion time using FRS by 22% (avg=103.095 s, $\sigma=19.637$ s). On the other hand,

---

[5]Gavin Moore, "Immersive Entertainment. Augmented Virtual Reality," https://medium.com/@gavizoid16/immersive-entertainment-augmented-virtual-reality-fc3dd2037c20
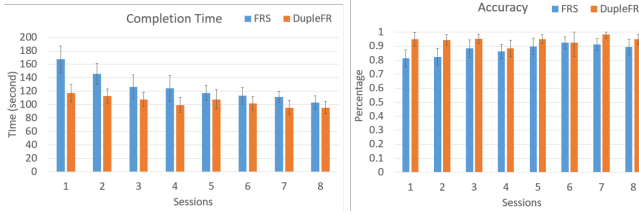
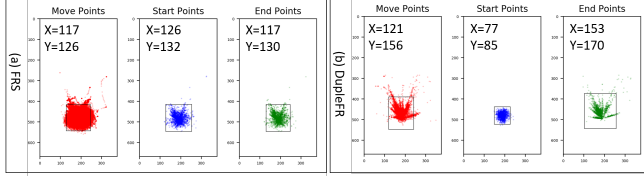Fig. 10. Completion Time (L) & Accuracy Rate (R) of FRS and DupleFR.

Fig. 11. Interaction footprint for (a) FRS and (b) DupleFR, where dots in red, blue and green colors represent the overall move points, start points and end points; Black boxes show the central interaction areas; X and Y values are the horizontal and vertical dimensions of the interaction areas.

DupleFR shows only a slight learning effect, with the participants improving by 9% over the 8 sessions (avg=95.498 s, σ=18.144 s). Equivalently, FRS and DupleFR demonstrate the completion time (per text acquisition) of 14.009 and 11.633 seconds across all trials as well as the completion time of 11.455 and 10.611 seconds in the final session.

2) Accuracy: Two-way repeated measures ANOVA demonstrates a significant effect of the Interaction Approaches and the Sessions ($F_{1,7}$ = 33.98, p <0.05), which indicates the significance of interaction approaches on the accuracy rate and the learning effect between sessions on similar interaction approach conditions. Figure 10 (right) shows the average accuracy rate for FRS and DupleFR over the 8 sessions. The participants with DupleFR (avg=94.26%, σ=0.1029) complete the tasks 6.57% more accurate than FRS (avg=87.69%, σ=0.1142). During the first session, FRS and DupleFR achieve the accuracy rate of 81.31% (σ=0.1206) and 89.39%. At the seventh session, participants with FRS and DupleFR reaches the highest accuracy rates of 91.41% (σ=0.0835) and 94.95% (σ=0.0953) respectively.

3) Footprint: Figure 11 shows the interaction footprint for FRS and DupleFR on an iPhone 7 with a 375 pt * 668 pt screen. To improve the readability, the red, green and blue dots shown in Figure 4 are footprint captured every second, while we compute the interaction areas (black box - the same approach as described in Section III) using samples captured every 50ms. As shown in the Move Points (Red dots) in Figure 11, FRS and DupleFR generate interaction areas of 117 pt * 126 pt and 121 pt * 156 pt, respectively. FRS interaction area is 25.24% smaller than DupleFR. We compute the distances between the centres of start points and end points. FRS and DupleFR display distances of 1.3232 pt and 23.7690 pt. The gesture design leads to a significant difference in the interaction areas as well as distances. FRS involves three gesture techniques that are ForceRay (FR), ForceSelect (FS) and circular gestures but the latter two techniques make a denser and more coherent pattern: (1) The participants with
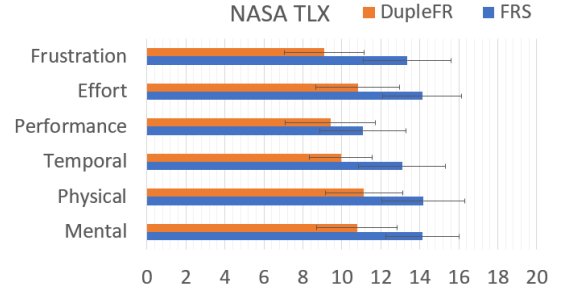

Fig. 12. NASA TLX results for FRS and DupleFR.

FS only hold their fingers on the fixed touchpoints; (2) The participants applying circular gestures limit their interaction near the button edge. Thus, a less deviated interaction area results from FS and circular gestures counteracting the initial step of FR. On the other hand, FR triggers larger interaction areas than the above two techniques. As DupleFR is composed of two successive FR, it displays a larger dispersion area. The radial pattern of DupleFR reflects the true space requirements supporting FR for the two interaction approaches. The end points (green dots) of DupleFR has the biggest dispersion among the six footprints (153 pt * 170 pt, equivalent to 12 * 13 mm^2 on the iPhone 7 screen of 326 ppi) in Figure 11. The end points of FR are regarded as a practical estimation of the interaction areas for the two interaction approaches, as FS and the circular gestures lead to underestimating of the average footprint. In a previous work [15], an area of 250 pt * 400 pt is measured because the participants can begin at random locations near the bottom bezel of the smartphone. In contrast, our study restricts our participants to start at a fixed and convergent location.

4) NASA TLX: We display the results of the NASA TLX survey in Figure 12. One-way ANOVA under two interaction approaches shows an effect of the interaction approach for all metrics (p <0.05) except for the Performance (p-value = 0.2243). These results imply that the users perceive a difference between the two systems in terms of Mental Demand (p=.0073), Physical Demand (p=.0181), Temporal Demand (p=.0106), Total Effort (p=.0118), and Frustration (p=.0020). In particular, the difference is most significant for Mental Demand and Frustration (p <0.01). Users found that FRS requires a significantly higher Mental Demand (1.68 points) than DupleFR, and its usage is much more frustrating (2.14 points). Interestingly, despite DupleFR showing an overall better completion time and a higher accuracy, users are unaware of a significant improvement in Performance.

C. Discussion

a) Performance: Both interaction approaches are modified from state-of-the-art techniques: ForceSelect (FS) achieves 8.5 seconds per text acquisition with direct manipulation on touchscreens [1], while ForceRay (FR) has an average of 1.672 seconds per reach to either a large or small icon [15]. Our approaches were tested with 9 sub-tasks in every session and hence FRS and DupleFR demonstrate on average 14.009

and 11.633 seconds for each text acquisition, respectively. DupleFR applies two ForceRays for each sub-task. Reaching for a character thus takes approximately 5.8165 seconds, where the two confirmation taps are neglected. It is important to note that the original implementation of ForceRay was targeting large icons, while we apply it to dense and small text content, leading to significantly slower completion times. In addition, the time sum of FS [1] and FR (our study) are close to FRS (14.009 seconds). Also, the completion time of [1] includes pan gestures that decide the number the textual units. Instead, we consider the circular gesture as an effective alternative technique to achieve similar time performance. Among the two approaches, DupleFR is significantly faster and more accurate than FRS. The key reason is that more lengthy procedures can be more error-prone and time-consuming. The participants also reflected that DupleFR is more intuitive than FRS. DupleFR needs less visual demand for selecting the two ends of the textual contents. The participants with FRS constantly needs visual demand to adjust the circular gesture for the designated textual units. Even though FS is regarded as the state-of-the-art technique on touchscreen interfaces, the participants on the non-touchable smartglasses prefer the solely FR-driven DupleFR interface to FRS involving FS.

b) Comparison with other modalities: As shown in the direct manipulation approaches designated to touchscreens [41], selection time with commercial standards (10.5 - 18.2 seconds) with accuracy rates ranged from 97.5% to 99.3%, and a state-of-the-art technique named BezelCopy (6.0 - 10.1 seconds) depends on the textual granularity with accuracy above 96.29%. After 8 training sessions, FRS and DupleFR show 11.455 and 10.611 seconds per textual acquisition, which are faster than commercial standards but slightly slower than the BezelCopy. However, FRS and DupleFR have lower average accuracy (87.69% and 94.26%), considering the disadvantage of indirect manipulation [44]. Our approaches are also more straightforward than speech commands. Voice interfaces cannot reach a target character or word in a precise manner, and recall operations are cumbersome [43]. Another voice interface reveals that users spent around 80.3% of their time in reviewing and editing the typed text and only 19.7% of the time in typing characters [42]. Furthermore, the commercial standards provide insufficient supports to textual granularity and hence obstruct further operations on textual contents such as copy-and-paste or translation of small text pieces. Users take images containing textual contents and potentially retype or extract the texts on the desktop environment, which deteriorates the mobility.

c) Interaction area: Even though DupleFR has a slightly larger interaction area than FRS, we still take the most dispersed area (153 pt * 170 pt / 12 * 13 mm^2) as the suggested interaction areas for two approaches as ForceRay is the most space-demanding technique supporting the two approaches. The thumb-size interface of FRS and DupleFR enables subtle interaction for higher social acceptability [27][28]. For instance, users may hide their smartphone in their pocket and perform unnoticeable thumb movements. This miniature-size interfaces can be developed into many form factors on smart wearables. For example, users can press on the vertical side of the spectacles frames of smartglasses. Alternatively, users can perform either thumb-to-finger interactions on smart rings or finger-to-arm interactions on smart wristbands.

d) Limitations and applications: This paper primarily focuses on the performance and space requirements of interaction approaches. However, the experiments were conducted on a smartphone touchscreen. The user performance and their perceptions on the aforementioned form factors of wearable devices are not yet investigated. In addition, the participants reported that they prefer to acquire long text with FRS more than DupleFR. Thus, the effects of text length on the interaction approaches should be investigated. A real-life application for managing textual contents in mobile scenarios can be further developed from the implemented interaction approaches. After taking a photo of the text material, Optical Character Recognition (OCR) can extract the text on the material. Next, the user with either FRS or DupleFR, manipulates the extracted text contents at various granularities. Finally, the system can obtain the contents in text format for further uses such as copy-and-paste operations. The commercial standard only allows users to take a photo and retype the content manually in the absence of directly reachable interfaces on mobile headsets.

VI. Conclusion

This paper serves as a groundwork to evaluate the feasibility of force-assisted interaction for acquiring text contents on mobile headsets. We studied three reachability techniques for textual content, representing a commercial standard and two state-of-the-art solutions. Among these techniques, ForceRay was significantly saving on-screen space, compared to Keyboard Trackpad and BezelCursor. Force-assisted techniques can effectively shrink the size of the interaction area. Accordingly, we designed and implemented two force-driven interaction approaches named ForceRaynSelect (FRS) and DupleFR for text acquisition on smartglasses. In our user study, participants managed textual contents at various granularities ranging from character to the entire text with both approaches. In the final session, participants using FRS and DupleFR respectively achieved average completion times of 103.095 and 95.498 seconds with accuracy rates of 91.41% and 94.95%. A miniature force-sensitive area (12 * 13 mm^2) can accommodate the two interaction methods that potentially improve the input bandwidth of wearable computers. The existing implementation on smartphones allows users to maintain a firm grip and perform one-thumb gestures in concealed manners. The evaluation results offer remarkable design implications in regard to input options on small wearable devices.

REFERENCES

[1] Alix Goguey, Sylvain Malacria, and Carl Gutwin. 2018. Improving Discoverability and Expert Performance in Force-Sensitive Text Selection for Touch Devices with Mode Gauges. In Proc. of the 2018 CHI Conf.(CHI '18). ACM, NY, USA, 12 pages.

[2] Motoki Miura and Kenji Saisho. 2014. A Text Selection Technique Using Word Snapping. Procedia Comp. Sci. 35 (2014), 1644–51.

[3] Axel Antoine et al. 2017. ForceEdge: Controlling Autoscroll on Both Desktop and Mobile Computers Using the Force. In Proc. of the 2017 CHI Conf.(CHI '17). ACM, NY, USA, 3281–3292.

[4] Ken Hinckley, Xiaojun Bi, Michel Pahud, and Bill Buxton. 2012. Informal information gathering techniques for active reading. In Proc. of the SIGCHI Conf.(CHI '12). ACM, NY, USA, 1893–1896.

[5] Pekka Parhi, Amy K. Karlson, and Benjamin B. Bederson. 2006. Target size study for one-handed thumb use on small touchscreen devices. In Proc. of the 8th Conf. on MobileHCI. ACM, NY, USA, 203–210.

[6] Sylvain Malacria et al. Push-Edge and Slide-Edge: Scrolling by Pushing Against the Viewport Edge. In Proc. of the 33rd Annual ACM Conf.(CHI '15). ACM, NY, USA, 2773–2776.

[7] Lik-Hang LEE, Kit-Yung LAM, Tong LI, Tristan BRAUD, Xiang SU and Pan HUI. Quadmetric Optimized Thumb-to-Finger Interaction for Force Assisted One-Handed Text Entry on Mobile Headsets. Proc. of the ACM on Ubicomp/IMWUT.

[8] Christian Corsten et al. 2017. Release, Don't Wait!: Reliable Force Input Confirmation with Quick Release. In Proc. of the 2017 ACM Int. Conf. on ISS '17. ACM, NY, USA, 246–251.

[9] Zhen Li et al. 2019. HoloDoc: Enabling Mixed Reality Workspaces that Harness Physical and Digital Content. In Proc. of the 2019 CHI Conf.(CHI '19). ACM, NY, USA.

[10] Masakazu Higuchi and Takashi Komuro. 2015. Multi-finger AR Typing Interface for Mobile Devices Using High-Speed Hand Motion Recognition. In Proc. of the 33rd Annual ACM Conf. Extended Abstracts(CHI EA '15). ACM, NY, USA, 1235–1240.

[11] Atsushi Umakatsu et al. 2012. Pinch-n-paste: direct texture transfer interaction in augmented reality. In Proc. of the 18th ACM Symp. on Virtual reality software and technology (VRST '12). ACM, NY, USA, 185–186.

[12] H. Kim, J. Lee, H. Yeo, A. Quigley and W. Woo, "SWAG Demo: Smart Watch Assisted Gesture Interaction for Mixed Reality Head-Mounted Displays," 2018 IEEE Int. Symp. on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct), Munich, Germany, 2018, pp. 428–429.

[13] S. Hegde et al. "A Fingertip Gestural User Interface Without Depth Data for Mixed Reality Applications," 2018 IEEE Int. Symp. on ISMAR (Adjunct), Munich, Germany, 2018, pp. 395–396.

[14] N. Chaconas and T. Hllerer, "An Evaluation of Bimanual Gestures on the Microsoft HoloLens," 2018 IEEE Conf. on Virtual Reality and 3D User Interfaces (VR), Reutlingen, 2018, pp. 1–8.

[15] Christian Corsten, Marcel Lahaye, Jan Borchers, and Simon Voelker. 2019. ForceRay: Extending Thumb Reach via Force Input Stabilizes Device Grip for Mobile Touch Input. In Proc. of the 2019 CHI Conf.(CHI '19). ACM, NY, USA, 12 pages.

[16] Wing Ho Andy Li, Hongbo Fu, and Kening Zhu. 2016. BezelCursor: Bezel-Initiated Cursor for One-Handed Target Acquisition on Mobile Touch Screens. Int. J. Mob. Hum. Comp. Interact. 8, 1 (Jan 2016), 1-22.

[17] Youli Chang et al. Understanding Users' Touch Behavior on Large Mobile Touch-Screens and Assisted Targeting by Tilting Gesture, Proc. of the 33rd Annual ACM Conf. , Apr 2015, Seoul, Republic of Korea.

[18] Sunjun Kim et al. Interaction techniques for unreachable objects on the touchscreen, Proc. of the 24th Australian Computer-Human Interaction Conf., p.295-298, Nov 2012, Melbourne, Australia.

[19] Hsin-Ruey Tsai et al. MovingScreen: selecting hard-to-reach targets with automatic comfort zone calibration on mobile devices, Proc. of the 18th Int. Conf. on MobileHCI Adjunct, Sept 2016, Florence, Italy

[20] Huy Viet Le et al. Investigating Screen Shifting Techniques to Improve One-Handed Smartphone Usage, Proc. of the 9th Nordic Conf. on Human-Computer Interaction, Oct 2016, Gothenburg, Sweden

[21] Rachel Eardley , Anne Roudaut , Steve Gill , Stephen J. Thompson, Understanding Grip Shifts: How Form Factors Impact Hand Movements on Mobile Phones, Proc. of the 2017 CHI Conf. , May 06-11, 2017, Colorado, USA

[22] Tovi Grossman , Ravin Balakrishnan, The bubble cursor: enhancing target acquisition by dynamic resizing of the cursor's activation area, Proc. of the SIGCHI Conf. , April 02-07, 2005, Oregon, USA

[23] Lik-Hang Lee and Pan Hui, "Interaction Methods for Smart Glasses: A Survey," in IEEE Access, vol. 6, pp. 28712-28732, 2018.

[24] Lik-Hang Lee, Tristan Braud, Farshid Hassani Bijarbooneh, and Pan Hui. TiPoint: Detecting Fingertip for Mid-Air Interaction on Computational Resource Constrained Smartglasses. In Proc. of the 23rd ACM Annual Int. Symp. on ISWC'19, London, September 2019

[25] Lik-Hang Lee, Kit Yung Lam, Yui Pan Yau, Tristan Braud, and Pan Hui. HIBEY: Hide the Keyboard in Augmented Reality. In Proc. of IEEE Int. Conf. on Pervasive Computing and Communications (PerCom 2019), Kyoto Japan, March 2019

[26] Kit Yung Lam, Lik-Hang Lee, Tristan Braud, and Pan Hui. M2A: A Framework for Visualizing Information from Mobile Web to Mobile Augmented Reality. In Proc. of IEEE Int. Conf. on Pervasive Computing and Communications (PerCom 2019), Kyoto Japan, March 2019

[27] B. Kollee, S. Kratz, and A. Dunnigan, Exploring gestural interaction insmart spaces using head mounted devices with ego-centric sensing, in Proc. of the 2Nd ACM Symp. on Spatial User Interaction, ser. SUI, 2014, pp. 40–49.

[28] S. Li, A. Ashok, Y. Zhang, C. Xu, J. Lindqvist, and M. Gruteser, Whose move is it anyway? authenticating smart wearable devices using uniquehead movement patterns, 03 2016, pp. 1–9.

[29] Zhanpeng Huang, Weikai Li, and Pan Hui. 2015. Ubii: Towards Seamless Interaction Between Digital and Physical Worlds. In Proc. of the 23rd ACM Int. Conf. on Multimedia (MM 15). ACM, NY, USA, 341–350.

[30] T. Lee and T. Hollerer. 2007. Handy AR: Markerless Inspection of Augmented Reality Objects Using Fingertip Tracking. In2007 11th IEEE Int. Symp. on Wearable Computers. 8390.

[31] Stephen A. Brewster and Michael Hughes. 2009. Pressure-Based Text Entry for Mobile Devices. In Proc. of the 11th Int. Conf. on MobileHCI. ACM, NY, USA, Article 9, 4 pages.

[32] Mingyuan Zhong et al. 2018. ForceBoard: Subtle Text Entry Leveraging Pressure. In Proc. of the 2018 CHI Conf. on Human Factors in Computing Systems (CHI 18). ACM, NY, USA, Article 528.

[33] Sachi Mizobuchi et al. Making an impression:Force-controlled pen input for handheld devices. InCHI 05 Extended Abstracts , CHI EA 05,pages 16611664, NY, USA, 2005. ACM.

[34] Jibin Yin et al. 2010. Pen pressure control in trajectory-based interaction. Behav. Inf. Technol. 29, 2 (March 2010), 137-148.

[35] Apple Inc. 2018. iOS 11 UIKit API Reference. https://developer.apple.com/reference/uikit/uitouch. [Online; accessed 30 Aug 2019].

[36] Olivier Chapuis et al. 2009. DynaSpot: Speed-dependent Area Cursor. In Proc. of the SIGCHI Conf.(CHI 09). ACM, NY, USA, 13911400.

[37] Christian Corsten et al. 2017. Release, Don't Wait!: Reliable Force Input Confirmation with Quick Release. In Proc. of the 2017 ACM Int. Conf. on ISS '17. ACM, NY, USA, 246-251.

[38] Craig Stewart et al. 2010. Characteristics of Pressure-Based Input for Mobile Devices. In Proc. of the SIGCHI Conf.(CHI 10). ACM, NY, USA, 801810.

[39] N.A.R.C. Human Performance Research Group.(1999) NASA Task Load Index (TLX). [Online]. Available: https://humansystems.arc.nasa.gov/groups/ TLX/downloads/TLX.pdf

[40] Yuki Matsuura et al. 2019. Readability and legibility of fonts considering shakiness of head mounted displays. In Proc. of the 23rd ISWC '19. ACM, NY, USA, 150-159.

[41] Chen, C., Perrault, S.T., Zhao, S., Ooi, W.T. (2014). BezelCopy: an efficient cross-application copy-paste technique for touchscreen smartphones. AVI.

[42] Shiri Azenkot and Nicole B. Lee. 2013. Exploring the use of speech input by blind people on mobile devices. In Proc. of the 15th Int. ACM SIGACCESS ASSETS '13.

[43] Jonggi Hong and Leah Findlater. 2018. Identifying Speech Input Errors Through Audio-Only Interaction. In Proc. of the 2018 CHI Conf.

[44] Cheng-Yao Wang et al. 2015. PalmType: Using Palms as Keyboards for Smart Glasses. In Proc. of the 17th Int. Conf. on MobileHCI.