International Journal of Human–Computer Interaction

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/hihc20

Depth and Breadth of Pie Menus for Mid-air Gesture Interaction

Wenmin Li , Xueyi Wan , Yanwei Shi , Nailang Yao , Ci Wang & Zaifeng Gao

INTERNATIONAL JOURNAL OF HUMAN–COMPUTER INTERACTION
https://doi.org/10.1080/10447318.2020.1809245

Depth and Breadth of Pie Menus for Mid-air Gesture Interaction

Wenmin Li, Xueyi Wan, Yanwei Shi, Nailang Yao, Ci Wang, and Zaifeng Gao

Department of Psychology and Behavioral Sciences, Zhejiang University, Hangzhou, China

ABSTRACT
Mid-air gesture interaction has become one of the most promising human–computer interaction manners. However, the principles of designing pie menus for mid-air gesture interaction are largely lacking. Currently, no study has addressed the breadth and depth of the pie menu for mid-air gesture interaction. Here, in two experiments, we found that the breadth and depth of a pie menu had distinct influences on the operation performance in mid-air gestural interaction: breadth affects both the accuracy and reaction time of the operation, while depth only affects reaction time. Overall, when operation accuracy is the main concern, the breadth of the pie menu will be the key factor, and it should not exceed eight items (Experiments 1 and 2). However, when the operation also emphasizes efficiency, the depth of the pie menu should be considered: fewer layers lead to faster responses, with two layers yielding the best performance (Experiment 2).

1. Introduction

With advances in human–computer interaction technologies such as multi-touch, speech recognition, gesture interaction, and gaze control, emerging natural user interfaces (NUI) are changing the way that users rely on control devices in graphical user interfaces (GUI). Among various new methods of interaction in NUI, gesture interaction has become one of the most promising fields (e.g., Chen et al., 2018; Pereira et al., 2015; Zhao et al., 2014). Gesture interaction enables users to interact with a device in an intuitive and easy manner without traditional input devices such as mouse or keyboard (e.g., Baudel & Beaudouin-Lafon, 1993; Chen et al., 2018; Davis et al., 2016; M. Lee et al., 2020; Norman, 2010; Pang et al., 2014). Compared with conventional interaction methods, gesture interaction is more natural for users (e.g., Baudel & Beaudouin-Lafon, 1993; Bolt, 1980; Chen et al., 2018; Pereira et al., 2015), and can improve users' engagement and emotional experience (Bianchi-berthouze et al., 2007). Therefore, gesture interaction shows great potential application for smart home, virtual games and intelligent driving (BMW Group, 2019; Kang et al., 2013; Liang, 2013). At the current stage, there are two typical categories of gestural interaction in general. In the first category, users physically touch an equipment (e.g., hand-held wands, hand-shape gloves) when performing a gesture interaction (e.g., Cao & Balakrishnan, 2003; M. Lee et al., 2020; Wilson & Shafer, 2003). In the second category, users fulfill the interaction by using mid-air gestures without touching any equipment. This mid-air gesture interaction is achieved by using body gestures in 3D space with the help of cameras (e.g., Kinect, Leap Motion; Brand et al., 2016; Ferron et al., 2019; Koutsabasis & Vogiatzidakis, 2019; M. Lee et al., 2020). Mid-air gestures can even be learned effectively via cross-modal training (e.g., Henderson et al., 2019). This study focused on mid-air gesture interaction.

So far, two distinct types of mid-air gesture interaction have emerged, according to whether a visual feedback (e.g., a cursor) is displayed to the user during the interaction (M. Lee et al., 2020). For the mid-air interaction without visual feedback, it is critical to design a set of user-cantered gestures to enhance interaction experience, which has been examined extensively (e.g., Chen et al., 2018; Choi et al., 2012; Pereira et al., 2015). However, for the mid-air gesture interaction with a visual feedback, although gestures are needed in a task, the feedback-related issues of gesture interaction become an important issue. For instance, in a cursor-based gesture interaction where a cursor is displayed to the users, users can use simple hand gestures (e.g., grabbing, dwelling) to perform the task, since the NUI system will track the positions of hand and map it to the menu or icon (e.g., Brand et al., 2016; M. Lee et al., 2020; Walter et al., 2014). In this case, since the users had to select an item from a menu, the design of the visual menu will considerably affect the performance and the experience of the mid-air gesture interaction.

Compatibility between the interaction manner and the visual menu is a golden standard in selecting an appropriate visual menu (G. Bailly et al., 2011; Gebhardt et al., 2013; Tian et al., 2007; Trinh et al., 2014). Currently, there are two main categories of visual menu: linear menu, and pie menu (or radial/marking menu; G. Bailly et al., 2017, 2011; Komerska & Ware, 2004; Urbina et al., 2010). The options in a linear menu are organized vertically or horizontally, while the options in a pie menu are arranged in a circular format (e.g., G. Bailly et al., 2017; Callahan et al., 1988; Urbina et al., 2010). A pie menu allows users to select a menu item by popping up a radial menu and moving the cursor from the

CONTACT Zaifeng Gao zaifengg@zju.edu.cn Department of Psychology and Behavioral Sciences, Zhejiang University, Hangzhou 310007, P.R. China

center to the desired item, or drawing a mark along the selecting path (e.g., G. Bailly et al., 2017; Kurtenbach, 1993; Kurtenbach & Buxton, 1993; Kurtenbach et al., 1993). Linear menus have been commonly employed in GUI owing to their interaction advantages (G. Bailly et al., 2017). In contrast to GUIs, mid-air gesture interaction has low operation precision (Norman, 2010) and tends to suffer from physical fatigue (Ferron et al., 2019; Kim et al., 2011). Befitting the characteristics of gesture interaction, pie menus have become a design interest due to its three inherent advantages. First, the target item can be selected by crossing the outer border of the menu without waiting for any dwell time (Komerska & Ware, 2004; Urbina et al., 2010), which can facilitate the establishment of operation memory (Appert & Zhai, 2009; G. Bailly et al., 2010; Callahan et al., 1988). Second, compared with linear menus, pie menus significantly reduce the moving distance of the body (G. Bailly et al., 2011; Callahan et al., 1988; Fitts, 1954), which is beneficial to reducing users' physical fatigue. Third, the direction of operation represents users' intention, and the area for displaying an option is relatively large. These two aspects help compromise the impact of low operation precision in gesture interaction. Indeed, Komerska and Ware (2004) compared linear and pie menus for gesture interaction in virtual reality, showing that selection via a pie menu is considerably faster and more accurate. Moreover, Leap motion, a popular product in the gesture interaction field, has extensively adopted pie menus for interaction.

Currently, research into the design of pie menus for gesture interaction is at its initial stages. Previous studies had explored the advantages of pie menus (e.g., Komerska & Ware, 2004; Mo & Zhou, 2018), methods of option selection (e.g., Komerska & Ware, 2004; Lim & Kristensson, 2019; Urbina et al., 2010; Walter et al., 2014), and spatial layout of pie menus (e.g., Davis et al., 2016; Gebhardt et al., 2013) in gesture interaction. It is of note that there had been studies focusing on invisible pie menus (i.e., marking menu; e.g., Kurtenbach & Buxton, 1993; Kurtenbach et al., 1993; Tapia & Kurtenbach, 1995). However, invisible pie menu stresses recall over recognition, and is suggested to be suitable for expert users (Henderson et al., 2020; Kurtenbach et al., 1993). Consequently, visible pie menu is currently recommended and is the type of pie menu discussed in this study. A normal menu of mid-air gesture interaction usually contains many operation options. Arranging these options requires us to consider issues of breadth and depth, which considerably affect the interaction experience (e.g., Miller, 1981; Seppala & Salvendy, 1985). The breadth refers to the number of options at one layer, and the depth refers to the number of levels when options are layered (Miller, 1981; Shneiderman, 2010). The breadth of a menu positively correlates with users' search time of the menu, and deeper depth of a menu increases users' working memory load (i.e., to remember the exact path to reach the target item, Commarford et al., 2008; Jacko & Salvendy, 1996) which leads to decreased accuracy and increased reaction time. When designing a visual menu, one should also consider the trade-off between breadth and depth, since an optimal menu needs to consider the combination of the two orthogonal dimensions (i.e., menu structure, Burnett et al., 2013; Kiger, 1984; Miller, 1981). However, to the best of our knowledge, no study has addressed these issues under the circumstance of mid-air gesture interaction.

The issues of breadth and depth have been extensively explored for linear menus in GUI (e.g., Eric Lee et al., 1985; Kurtenbach & Buxton, 1993; Miller, 1981; Paap & Roske-Hofstrand, 1986; Snowberry et al., 1983), offering guiding principles on menu design. For example, the optimal menu breadth is approximately 4–8 items (Eric Lee et al., 1985), and the breadth should not exceed 10 for unstructured menus (Paap & Roske-Hofstrand, 1986); the optimal depth is 2 layers and no more than three in hierarchical menus (e.g., Kiger, 1984; Mo & Zhou, 2018); a menu that exceeds these limits would dramatically impair users' performance. However, these conclusions for GUI cannot be directly applied to the design of pie menus, especially in mid-air gesture interaction, because the interaction characteristics in gesture interaction are distinct from those of GUI (e.g., Norman, 2010; Pang et al., 2014; Zhao et al., 2014). First, mid-air gesture interaction is designed for low-precision task (e.g., virtual game and smart home), while GUI is suitable for high-precision task (i.e., document and design). Menus in gesture interaction usually contains a limited number of items, and are controlled directly by users' hands or arms, while menus in GUI usually contain several items, and are controlled indirectly by a mouse. This may lead to different limits and different combinations of menu breadth and depth. Second, mid-air gesture interaction is operated directly by hands over a large spatial scope, and thus is prone to cause fatigue, which implies that pie menu in gestural interaction may be more sensitive to changes in breadth and depth. Thus, a different limit and breadth–depth combination may emerge for gestural interaction. To date, in the field of NUI, only one study has considered the breadth and depth of pie menus under a circumstance of gaze control interaction (Urbina et al., 2010). Urbina et al. compared four breadth levels (4, 6, 8, and 12), and found that an optimal menu should contain no more than six options in one layer. Additionally, they manipulated three levels of depth (2, 3, and 4), revealing a non-significant impact of depth on interaction performance. It is worth noting that Urbina et al. (2010) explored the two issues without considering the trade-off between breadth and depth, hence the distinct numbers of total items may contaminate their findings. Moreover, mid-air gesture interaction, in essence, is different from the interaction of gaze control. Gaze control is difficult to control the involuntary eye movement and drift, while hand and arm gestures are easy to control but are more prone to physical fatigue. Empirical evidence is required to directly elucidate the breadth and depth for mid-air gesture interaction.

To this end, we investigated the depth and breadth of pie menus in mid-air gesture interaction, while considering the trade-off between them. As Leap Motion is one of the most popular controllers in mid-air gesture interaction (Bachmann et al., 2018), we adopted it as the equipment for hand gesture recognition in the experiments. Considering that selecting an item via crossing the outer border of the menu has been revealed as a better and promising gesture (e.g., Huckauf & Urbina, 2008; Urbina et al., 2010) and has been used in applications of Leap

Motion, we decided to explore the current issues by using this manner to select the target. Experiment 1 investigated the limit of breadth in a pie menu, and Experiment 2 further explored the limit of depth, taking the trade-off of depth and breadth into consideration.

2. Experiment 1: Breadth limit of pie menus

We first identified the limits of breadth for a pie menu in a single layer. According to previous studies (e.g., Miller, 1981; Paap & Roske-Hofstrand, 1986; Urbina et al., 2010), the limit of menu breadth is approximately 8–10 items in GUI and 6 items in gaze control interaction. We determined five levels of breadth (4, 6, 8, 10, or 12) for mid-air gesture interaction (see also Urbina et al., 2010 for a similar range), such that we could systematically check how the performance was affected by the breadth of a pie menu and determined the limit of menu breadth.

2.1. Method

2.1.1. Participants

Twenty-four participants (12 males) were recruited, with an average of 21.38 (SD = 1.60) years old. All participants were students at Zhejiang University, who were right-handed, had normal or corrected-to-normal vision acuity, and no physical disabilities. All participants did not have any interaction experience with Leap Motion.

2.1.2. Apparatus and stimuli

The pie menus were presented on a 23-inch Dell LED monitor with a resolution of 1920 × 1080 pixel at a refresh rate of 60 Hz. The participants were seated in a lighted room, interacting with the menus via the Leap Motion controller (LM-010, Leap Motion Inc.). The Controller was placed 22 and 28 cm away from the participant and screen, respectively. Twelve Chinese characters, chosen from the *Modern Chinese Frequency Dictionary* (Institute of Language Teaching, Beijing Language and Culture University, 1986), were used: 的(de), 是(shi), 这(zhe), 说(shuo), 和(he), 到(dao), 里(li), 来(lai), 把(ba), 看(kan), 要(yao), 很(hen). These characters (font Microsoft Yahei, font size 18) were used to simulate options in the menu,[1] with the strokes ranging from 7 to 9 and their usage frequency ranking the top 12 in daily life of China. The items in the menu of each trial were randomly selected from the 12 characters. The target item appeared evenly at each location (e.g., a target would be present twice at each location of the pie menu when the menu breadth was 12, since there were 24 trials in each breadth condition), while the positions of other non-target items were randomly selected from the left positions (i.e., menu breadth -1). All 12 characters were selected evenly as the target item. The pie menu had a radius of 100 pixels for the inner circle and 200 pixels for the outer circle. The visual distance between a menu option and the center of the pie was about 9° in visual angle.

The Leap Motion were used to track hand positions and two major gestures. A circular cursor was used as a feedback, representing the position of user's hand (see also Walter et al., 2014). For selection, we adopted *point* gesture, and the user was told to point at certain field and move in a specific direction to select items (Walter et al., 2014). For confirmation, we adopted the *crossing border* gesture, and the user was told to move the cursor to cross the outer border (like dragging out the target item, Huckauf & Urbina, 2008; Komerska & Ware, 2004; Urbina et al., 2010).

2.1.3. Design and procedure

The experiment adopted a one-factor (menu breadth: 4, 6, 8, 10, and 12) within-subjects design. Each breadth contained 24 trials, leading to a total of random120 trials. The experiment was then divided into four blocks, each block contained 30 trials. Participants had a 2-min rest after each block.

Participants first learned to interact with the menu via Leap Motion, until they could operate smoothly. This session lasted about 3 min. Then they practiced the experimental task for at least 10 trials. Once their overall accuracy was no less than 90%, the formal experiment initiated. The whole experiment lasted approximately 15 min.

Figure 1 shows the procedure used in Experiment 1. Each trial started with a gray circle (RGB: 110, 110, 110; radius: 30 pixels), containing a black target item. The target item turned to blue (RGB: 30, 144, 255) after 1 s, informing the participants to start interaction. Once Leap Motion detected the hand gesture, a green cursor (RGB: 160, 255, 160; radius: 20 pixels.) appeared to indicate the position of the operation. When the center of the green cursor crossed the boundary of the gray circle containing the target, the gray circle disappeared while the pie menu appearing at the screen center. To remind participants the to-be-selected item, the menu option was pre-activated when the distance between the center of the green cursor and the inner border of the option was less than 70 pixels. This was achieved by changing the color of the corresponding option slice into light gray (RGB: 200, 200, 200). Critically, when participants determined to select the item by moving the center of cursor across the inner border of the pie menu, the corresponding slice moved simultaneously with the cursor and was highlighted by changing color to dark gray (RGB: 100, 100, 100). Once the center of the cursor crossed the outer border of the menu, the menu option was successfully selected. Reaction time and accuracy were recorded. The reaction time was measured from the onset of the pie menu until participants selected an item in the menu; accuracy was determined according to whether the selected item matched the target item.

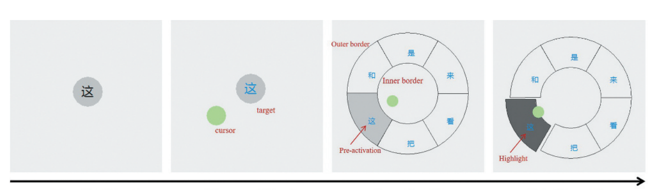

Figure 1. A schematic illustration of the experimental procedure in Experiment 1, when 6 items were presented. The English words in the figure were used for illustrating, and not displayed in the experiment.

2.1.4. Analysis

All trials were analyzed for accuracy, while only corrected trials were analyzed for reaction time. For accuracy and reaction time, we distinguished two types of responses: responding correctly without adjustment (one-shot response), and responding correctly regardless of adjustment (overall response).

A one-way repeated analysis of variance (ANOVA), by taking menu breath (4, 6, 8, 10, and 12 items) as the within-subjects factor, was conducted on both the accuracy and reaction time. For the significant main effect, post hoc contrasts (Bonferroni corrected) were conducted.

2.2. Results

For overall responses, the accuracy was above 90% for all the five breath levels, and the reaction time increased linearly as the breadth raised. Confirming the observation, the ANOVA on accuracy (Figure 2(a)) revealed a significant main effect of menu breadth, $F(4,92) = 7.26$, $p < .005$, $\eta_p^2 = .24$. Post hoc contrasts revealed significant differences between breadth 4 and 10 ($p < .005$), breadth 6 and 10 ($p < .05$), as well as breadth 4 and 12 ($p < .05$). Similarly, the ANOVA on reaction time (Figure 2(b)) revealed a significant main effect of menu breadth, $F(4,92) = 118.08$, $p < .001$, $\eta_p^2 = .843$. Post hoc contrasts revealed that the difference between each level of menu breadth was significant ($p < .005$).

For one-shot responses, the accuracy dropped linearly as the breath increased, and the accuracy was below 90% when the breadth beyond 8 items. Additionally, the reaction time increased linearly as the breadth raised. Confirming the observation, the ANOVA on accuracy (Figure 2(c)) revealed a significant main effect of menu breadth, $F(4,92) = 17.43$, $p < .005$, $\eta_p^2 = .43$. Post hoc contrasts revealed significant differences between breadth 4 and 8, breadth 4 and 10, breadth 4 and 12, breadth 6 and 10, breadth 6 and 12, as well as breadth 8 and 12 ($p < .05$). The ANOVA on reaction time (Figure 2(d)) revealed a significant main effect of menu breadth, $F(4,92) = 131.15$, $p < .001$, $\eta_p^2 = .85$. Post hoc contrasts revealed that the difference between each level of menu breadth was significant ($p < .005$).

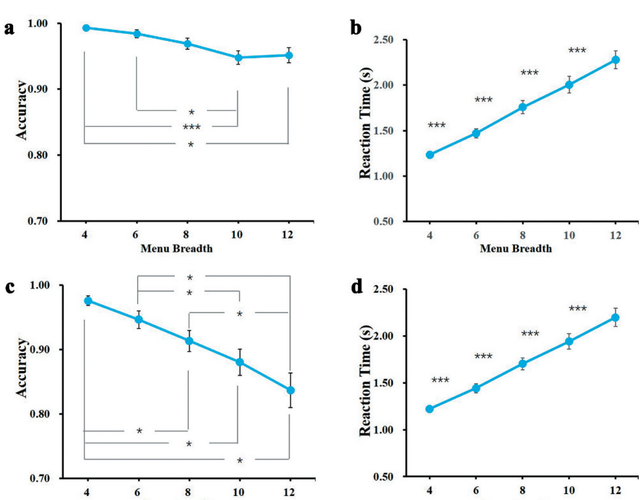

Figure 2. Results for Experiment 1. (a) The mean accuracy of overall response for different menu breadths. (b) The mean reaction time of overall response for different menu breadths. (c) The mean accuracy of one-shot response for different menu breadths. (d) The mean reaction time of one-shot response for different menu breadths. * $p < .05$; *** $p < .001$. Error bar stands for standard errors.

2.3. Discussion

Experiment 1 showed that the menu breath significantly affects both the accuracy and reaction time of users' operation. The more the items contained in a pie menu, the longer the reaction time required. Critically, the results on accuracy shed light on the upper limit of menu breadth. Particularly, accuracies of overall responses under 4 and 6 items were significantly better than that of 10 items, without significant differences relative to 8 items. Meanwhile, the accuracy of one-shot response was significantly higher than 90% for 4 ($p < .01$) and 6 ($p < .01$) items, with an averaged accuracy of 91% for 8 items, which was not significantly different from 90% ($p > .05$); the averaged accuracy continued dropping after 8 items. Moreover, the accuracy of one-shot responses for 4 items was significantly better than that of 8 items, without significant differences between 4 and 6 items, or between 6 and 8 items. Taking these results into consideration, we suggest the upper limit of the breadth of pie menus is approximately 8 items for mid-air gesture interaction.

3. Experiment 2: Depth limit in pie menus

Experiment 2 explored the depth limit of pie menus. We considered that it is important to keep the total number of menu items being constant. Only in this way, can we avoid the potential contamination of distinct numbers of menu items on the results and enable the recommended parameters to be useful in the real design activities (e.g., Burnett et al., 2013; Kiger, 1984; Miller, 1981). Since gesture interaction is prone to physical fatigue and fits an application context without too many menu options, we examined a case containing 16 options in total. Sixteen options enabled us to test four menu structures ($2 \times 2 \times 2 \times 2$, $2 \times 2 \times 4$, 2×8, 4×4),[2] which cover a range of depth values within the limit of breadth. The $2 \times 2 \times 2 \times 2$ had four layers (depth = 4), each containing two options; $2 \times 2 \times 4$ had three layers (depth = 3), with two options for the first and second layer and four options for the third layer. Furthermore, 2×8 had two layers (depth = 2), with two and eight options for the first and second layer, respectively; finally, 4×4 had two layers (depth = 2), each containing four options.

3.1. Method

3.1.1. Participants

Twenty-four new participants (10 males) were recruited for Experiment 2, with an average of 20.83 ($SD = 1.74$) years old. All participants were students at Zhejiang University, who were right-handed, had normal or corrected-to-normal vision acuity, and no physical disabilities. All participants did not have any interaction experience with Leap Motion.

3.1.2. Apparatus and stimuli

To simulate the hierarchy in the commands, we replaced Chinese characters in Experiment 1 with Chinese daily words. Particularly, two word lists were chosen from commonly used words in daily life for the item candidates in the pie menu (see Table 1). Each list had 16 items, which could be divided into four levels (for instance, plants [level 1], flowers & grass [level 2], flowers [level 3], rose [level 4]). The hierarchy of the pie menu was set according to the categorical relation of words. Additionally, when participants descend a menu hierarchy, if both parent menus and child submenus remain on the screen, layout overlap and visual crowding would appear (e.g., Gebhardt et al., 2013; Tapia & Kurtenbach, 1995). To avoid this problem, we hid the upper level menus and displayed the target submenu after the user selected it (cf. Tapia & Kurtenbach, 1995). To enable participants to be back to an upper level at will, we added a "back" option ("返回" in Chinese) in each submenu (cf. Davis et al., 2016). The back option always presented at the lower right corner of the pie menu (see Figure 3 for an illustration). The target item appeared evenly at each location except for the back option location, and the other items were randomly arranged. Other apparatus and stimuli were the same as those in Experiment 1.

3.1.3. Design and procedure

The experiment adopted a one-factor (menu structure: $2 \times 2 \times 2 \times 2$, $2 \times 2 \times 4$, 2×8, and 4×4) within-subjects design. The experiment was divided into four blocks according to the menu structure, the order of which was counterbalanced between participants using a Latin square manner. Each block had 16 trials, leading to a total of 64 trials. Participants had a 2-min rest after each block.

Participants first learned to interact with the menu via Leap Motion, until they could operate smoothly (familiar

Table 1. Item candidates used in the pie menu of Experiment 2. Chinese words inside the parentheses, instead of English words, were presented during the experiment.

Material	First level	Second level	Third level	Target item (Fourth level)
List A	Plants (植物)	Flowers & Grass (花草)	Flowers (花类)	White Flower (白花)
				Rose (玫瑰)
			Grass (草类)	Water Plant (水草)
				Asparagus Fern (文竹)
		Fruits & Vegetables (果蔬)	Fruits (水果)	Watermelon (西瓜)
				Apple (苹果)
			Vegetables (蔬菜)	Chives (韭菜)
				Spinach (菠菜)
	Animals (动物)	Birds & Beasts (鸟兽)	Birds (鸟类)	Sparrow (麻雀)
				Swallow (燕子)
			Beasts (兽类)	Lions (狮子)
				Tiger (老虎)
		Insects & Fishes (虫鱼)	Insects (昆虫)	Dragonfly (蜻蜓)
				Bee (蜜蜂)
			Fishes (鱼类)	Carp (鲤鱼)
				Crucian (鲫鱼)
List B	Clothing & Food (衣食)	Food & Drink (饮食)	Food (食品)	Beef (牛肉)
				Bok choy (青菜)
			Drink (饮品)	Cola (可乐)
				Black Tea (红茶)
		Dressing (穿着)	Clothes (衣服)	Sweater (毛衣)
				Trousers (长裤)
			Shoes & Socks (鞋袜)	Leather Shoe (皮鞋)
				Sock (短袜)
	Housing & Transportation (住行)	Housing (住房)	Resting (休息)	Bedroom (卧室)
				Living Room (客厅)
			Wash supplies (洗漱)	Bathroom浴室
				Pool (水池)
		Transportation (出行)	Self-driving (自驾)	Saloon Car (轿车)
				Motorcycle (摩托)
			Public (公共)	High-speed Train (高铁)
				Bus (公交)

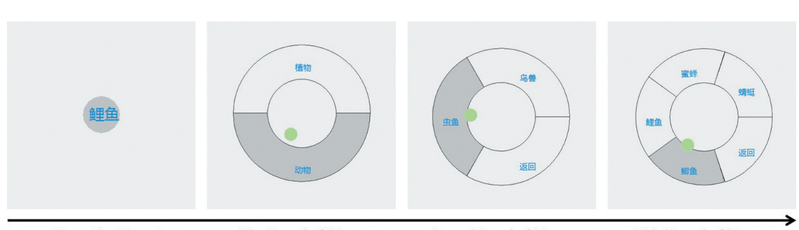

Figure 3. A schematic illustration of the experimental procedure in Experiment 2. Here shows an example of a $2 \times 2 \times 4$ menu structure.

stage). All the participants fulfilled the familiar stage within 3 min. Then they moved to the next stage, practicing the experimental task for at least 10 trials (practice stage). Once their overall accuracy was no less than 80%, the formal experiment initiated (formal stage). The whole experiment lasted approximately 15 min. It is worth noting that half of participants used list A in the familiar stage, and list B in the practice and formal stages, and a reversed order was used for the other half of participants.

Figure 3 shows the general procedure used in Experiment 2. Other aspects were the same as Experiment 1, except for the display of the pie menu, which could contain two, three, or four hierarchical layers. Particularly, since the target always displayed at the last layer, participants had to search the pie menu across different levels of items until locating the target. For menu structure 2 × 2 × 2 × 2, participants were presented the items at the first, second, third, and fourth levels of Table 1 in the four layers, respectively; for menu structure 2 × 2 × 4, participants were presented the items at the first, second, and fourth levels of Table 1 in the three layers, respectively; for menu structure 2 × 8, participants were presented the items at the first and fourth levels of Table 1 at the two layers, respectively; for the menu structure 4 × 4, participants were presented the items at the second and fourth levels of Table 1 in the two layers, respectively. For instance, if the target item was Carp for the menu structure 2 × 2 × 2 × 2, participants had to select Animals (动物), Insects & Fishes (虫鱼), Fishes (鱼类), and Carp (鲤鱼) sequentially. Task reaction time, accuracy, and number of returns were recorded. The reaction time was measured from the onset of the pie menu until participants selected one item in the menu; accuracy was determined according to whether the selected item matched the target item. The number of returns accumulated when participants selected the back option to return to a higher level in the hierarchical pie menu, which indicated that participants considered that they had taken a wrong path (Mo & Zhou, 2018).

3.1.4. Analysis

All trials were analyzed for accuracy, while only corrected trials were analyzed for reaction time. For both accuracy and reaction time, we distinguished two types of responses: responding correctly without adjustment (one-shot response), and responding correctly regardless of adjustment (overall response).

A one-way repeated ANOVA, by taking menu structure (2 × 2 × 2 × 2, 2 × 2 × 4, 2 × 8, 4 × 4) as the within-subjects factor, was conducted on the accuracy, reaction time and number of returns. For the significant main effect, post hoc contrasts (Bonferroni corrected) were conducted.

3.2. Results

For overall responses, the accuracy was above 90% for all the four structures, and the reaction time dropped linearly from four to two layers. The ANOVA on accuracy (Figure 4(a)) revealed a significant main effect of menu structure, $F(3,69) = 4.17$, $p < .05$, $\eta_p^2 = .15$. Post hoc contrasts did not reveal any

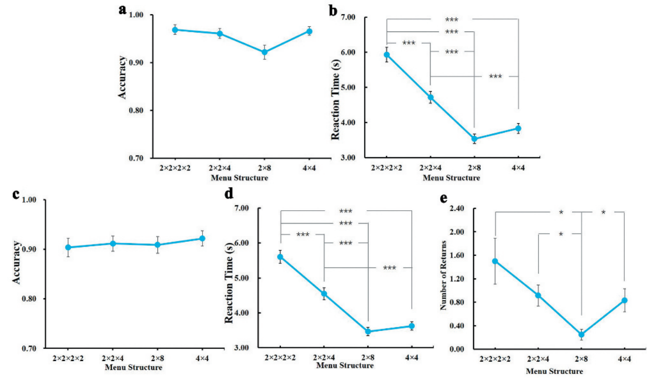

Figure 4. Results for Experiment 2. (a) The mean accuracy of overall response for different menu structures. (b) The mean reaction time of overall response for different menu structures. (c) The mean accuracy of one-shot response for different menu structures. (d) The mean reaction time of one-shot response for different menu structures. (e) The mean number of returns for different menu structures. * p <.05; *** p <.001. Error bar stands for standard errors.

significant difference between the four structures. The ANOVA on reaction time (Figure 4(b)) revealed a significant main effect of menu structure, $F(3,69) = 85.38$, $p < .005$, $\eta_p^2 = .79$. Post hoc contrasts revealed that the reaction time between any two menu structures reached significance ($p < .001$) except for the difference between menu structure 2 × 8 and 4 × 4 ($p > .05$).

For one-shot responses, the accuracy was around 90% for all the four structures, and the reaction time dropped linearly from four to two layers. The ANOVA on accuracy (Figure 4(c)) did not reveal a significant main effect of menu structure, $F(3,69) = .25$, $p = .86$, $\eta_p^2 = .01$. However, the ANOVA on reaction time (Figure 4(d)) revealed a significant main effect of menu structure, $F(3,69) = 101.48$, $p < .001$, $\eta_p^2 = .82$. Post hoc contrasts demonstrated that the reaction time between any two menu structures reached significance ($p < .001$) except for the difference between menu structure 2 × 8 and 4 × 4 ($p > .05$).

Finally, for the number of returns (Figure 4(e)), the menu structure 2 × 8 ($M = 0.25$) exhibited the least number of returns. Confirming this observation, the ANOVA revealed a significant main effect of menu structure, $F(3,69) = 5.12$, $p < .05$, $\eta_p^2 = .18$. Post hoc contrasts revealed significant differences ($p < .05$) between menu structure 2 × 8 and 2 × 2 × 2 × 2, menu structure 2 × 8 and 2 ×2 × 4, as well as menu structure 2 × 8 and 4 × 4.

3.3. Discussion

In Experiment 2, we found that the menu structure did not exert a dramatic impact on the response accuracy; however, it had a considerable influence on reaction time and number of returns, which enabled us to deduce the optimal menu structure and optimal depth. Particularly, we found that menu structures 2 × 8 and 4 × 4 (i.e., depth = 2) showed faster reaction time and there was no significant difference between them. The reaction time increased significantly from two layers (i.e., 2 × 8, 4 × 4) to three layers (2 × 2 × 4) and to four layers (2 × 2 × 2 × 2). These results indicated that the optimal depth was 2 and reaction time were more sensitive to menu depth than to breadth. Additionally, we found that even though depth was the same for the menu structures 2 × 8 and 4 × 4, menu structure 2 × 8 exhibited the fewest number of returns and fastest reaction time, suggesting that the optimal menu structure is 2 × 8 under the circumstance of 16 items, and the optimal depth is 2. This finding may be due to the fact that too many options in the first layer may raise user's uncertainty about the category that the target item belongs to. From this perspective, this finding is consistent with the view that a two-layered menu with larger breadth on the second layer leads to better performance (e.g., Norman & Chin, 1988).

4. General discussion

In the field of gestural interaction, pie menus are gaining considerable attention and becoming one of the main interactive menu types. However, the principles of designing pie menus for mid-air gesture interaction are largely lacking. This study, for the first time, explored the breadth and depth of pie menu in mid-air gesture interaction. We found that the breadth and depth of a pie menu have distinct influences on users' performance: breadth affects both the accuracy and reaction time of the operation, while depth only affected reaction time. The current findings shed light on the guidelines to the design of pie menus of mid-air gesture interaction. Specifically, if operation accuracy is the main concern, the breadth of the pie menu is the key factor in designing the pie menu, the upper limit of which is approximately eight items (Experiments 1 and 2). However, if the application context requires efficient interaction, the depth of the pie menu should also be considered: fewer layers lead to faster responses, with two layers yielding the best performance (Experiment 2). It is worth noting that the upper breadth limit falls within Miller's (1956) estimation of short-term memory capacity. Considering the trade-off between breadth and depth (with fixed number of menu items), the pie menu could take advantage of such limitation and thereby reduces the depth.

This study contributes to the pie menu design in NUI. Previously, only Urbina et al. (2010) addressed the depth and breadth of pie menus for gaze control interaction. This study addressed this issue in a new interaction manner. Moreover, this study is the first to consider the trade-off between breadth and depth when examining the issue of depth in NUI. Relative to the findings of Urbina et al. (2010), the upper limit of the breadth (eight) in gesture interaction is slightly larger than that in gaze control interaction (six), perhaps because it is harder to accurately control eye movements (due to tremor and drift) than body movements. On the other hand, in line with Urbina et al. (2010), we did not find an effect of depth in term of accuracy. While this finding may be due to that only three levels of depth (2, 3, and 4) were examined in both studies, it implies that breadth, rather than depth, of a pie menu exerts a larger impact over users' performance in NUI. Moreover, both Urbina et al. (2010) and this study found that the depth of a pie menu affects the reaction time, with the fastest responses at two layers; moreover, this study further demonstrated that more layers resulted in more return operations. These results suggest that designers should pay attention to depth when considering the interacting efficiency in NUI. In pie menu during gesture interaction, minimizing the depth could significantly reduce the complexity of menu structures (i.e., the complexity of paths to target options, Jacko & Salvendy, 1996). Taking this study and Urbina et al. (2010) together, we argue that for NUI, the maximum breadth in a pie menu should not exceed eight items, and the optimal depth is two layers. When menu items must exceed both the limits, expansion in breadth is recommended over expansion in depth.

Interestingly, although GUI and NUI have distinct processing mechanisms, the aforementioned recommendations are largely in line with those reported in pie menu studies on GUI, which indicated that the maximum breadth of the pie menu is 4–8 and the optimal depth is generally 2 layers (e.g., Gran, 2018; Kurtenbach & Buxton, 1993; Mo & Zhou, 2018; Zhao & Balakrishnan, 2004). These facts implied that the low precision of users' operation and physical fatigue embedded in gesture interaction do not significantly influence the depth and breadth limit of pie menus. Therefore, a general design principle might exist for designing pie menus across GUI and NUI: maximum breadth is eight items and the optimal depth is two layers.

Finally, it is of note that this study only explored the depth of pie menus with 16 items, future study may consider verifying the conclusion with a menu containing 27 items (menu structure: 3 × 3 × 3, 3 × 9), or 64 items (e.g., menu structure: 2 × 2 × 2 × 2 × 4, 4 × 4 × 4, 8 × 8), with a Kinect device (controlling by hands and whole-body movements). It is also worth exploring variants of pie menu layout, for instance, half pie layout which are popular in VR applications and are suitable for right-handed operations (G. Bailly et al., 2017; Hesselmann et al., 2009). Moreover, it has been found that expert users have different search patterns from novice users: Novices spend more time searching for individual items while experts spend more time on the motor aspects of acquisition (Ahlström et al., 2010). Therefore, whether experience with the GUI devices affects the layout of the pie menu also deserves attention in future. Related to this experience factor, participants may have already learned a location in the pie menu for the item in a trial in this study. However, because the menu items were the same despite changing orientations, this setting may lead to a potential confusion for participants during their operation in a following trial, which may slightly affect the result pattern reported here. Future study may consider using distinct menu items in each trial to verify the findings of this study.

5. Conclusion

This study investigated the breadth and depth of pie menu in mid-air gesture interaction. According to the results of two experiments, we reached two main findings: (1) The maximum breadth in a pie menu of mid-air gesture interaction should be within eight items, and the optimal depth is two layers; (2) when menu items had to exceed both limits, expansion in breadth is recommended over expansion in depth.

Moreover, the current findings imply a general design principle for designing pie menus across GUI and NUI.

Notes

1. It could be argued that the characters/words used in current study were chosen from daily life rather than from commands for gesture interaction (e.g., copy and paste), and thus lacked ecological validity. In fact, commonly used words and commands should not significantly affect the results. Corroborating this view, most previous studies on menu design adopted common words instead of real commands (e.g., Kurtenbach & Buxton, 1993; Miller, 1981; Urbina et al., 2010).
2. Studies (e.g., Norman & Chin, 1988) have indicated that a menu structure offers better performance when the breadth of the underlying layer is large. For instance, a menu structure 2 × 8, wherein the first and second layer contains two and eight items, respectively, is better than a menu structure 8 × 2, wherein the first and second layer contains eight and two items, respectively. Therefore, only one menu structure was explored for a depth–breadth combination (e.g., only menu structure 2 × 8 was explored).

Acknowledgments

This research was supported by Science and Technology Innovation 2030 program (2018AAA0101605), Fundamental Research Funds for the Central Universities National, Natural Science Foundation of China Grants (31771202), and Key Program of Natural Science Foundation of Zhejiang Province (LZ20C090001). We thanked Ziwei Liu, Jinnan Song, Ke Wang for the help in data collection.

Disclosure of potential conflict of interest

There is no financial interest or benefit that has arisen from the direct applications of your research.

References

Ahlström, D., Cockburn, A., Gutwin, C., & Irani, P. (2010, April). Why it's quick to be square: Modelling new and existing hierarchical menu designs. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1371–1380). Atlanta, Georgia, USA.

Appert, C., & Zhai, S. (2009, April). Using strokes as command shortcuts: Cognitive benefits and toolkit support. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 2289–2298). Boston, Massachusetts, USA: ACM.

Bachmann, D., Weichert, F., & Rinkenauer, G. (2018). Review of three-dimensional human-computer interaction with focus on the leap motion controller. *Sensors*, *18*(7), 2194. https://doi.org/10.3390/s18072194

Bailly, G., Lecolinet, E., & Guiard, Y. (2010, April). Finger-count & radial-stroke shortcuts: 2 techniques for augmenting linear menus on multi-touch surfaces. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 591–594). Atlanta, Georgia, USA: ACM.

Bailly, G., Lecolinet, E., & Nigay, L. (2017). Visual menu techniques. *ACM Computing Surveys (CSUR)*, *49*(4), 60. https://doi.org/10.1145/3002171

Bailly, G., Walter, R., Ller, J., Ning, T., & Lecolinet, E. (2011). Comparing free hand menu techniques for distant displays using linear, marking and finger-count menus. *Human-Computer Interaction - INTERACT 2011 -, Ifip Tc 13 International Conference, Lisbon, Portugal*, September 5-9, 2011, *Proceedings* (Vol.6947, pp.248-262). DBLP.

Baudel, T., & Beaudouin-Lafon, M. (1993). Charade: Remote control of objects using free-hand gestures. *Communications of the ACM*, *36*(7), 28–35. https://doi.org/10.1145/159544.159562

Bianchi-berthouze, N., Kim, W. W., & Patel, D. (2007). Does body movement engage you more in digital game play? and why? *Affective Computing and Intelligent Interaction, Second International Conference, ACII 2007, Lisbon, Portugal*, September 12-14, 2007, *Proceedings* (Vol.4738, pp.102-113). DBLP.

BMW Group. (2019, February) *Natural and fully multimodal interaction with the vehicle and its surroundings*. BMW Group PressClub. https://www.press.bmwgroup.com/usa/article/detail/T0292448EN_US/natural-and-fully-multimodal-interaction-with-the-vehicle-and-its-surroundings-bmw-group-presents-bmw-natural-interaction-for-the-first-time-at-mobile-world-congress-2019?language=en_US

Bolt, R. A. (1980, July). "Put-that-there" Voice and gesture at the graphics interface. *Proceedings of the 7th annual conference on Computer graphics and interactive techniques* (pp. 262–270). Cambridge, Massachusetts, USA.

Brand, D., Büchele, K., & Meschtscherjakov, A. (2016, October). Pointing at the HUD: Gesture interaction using a leap motion. *Adjunct Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 167–172). Ann Arbor, MI, USA.

Burnett, G. E., Lawson, G., Donkor, R., & Kuriyagawa, Y. (2013). Menu hierarchies for in-vehicle user-interfaces: Modelling the depth vs. breadth trade-off. *Displays*, *34*(4), 241–249. https://doi.org/10.1016/j.displa.2013.07.001

Callahan, J., Hopkins, D., Weiser, M., & Shneiderman, B. (1988). An empirical comparison of pie vs. linear menus. *CHI ' 88:Proc. SIGCHI conference on Human factors in computing systems* (pp.95–100). Washington D.C. USA: ACM Press.

Cao, X., & Balakrishnan, R. (2003, November). VisionWand: Interaction techniques for large displays using a passive wand tracked in 3D. In *Proceedings of the 16th annual ACM symposium on User interface software and technology* (pp. 173–182). Vancouver, BC, Canada.

Chen, Z., Ma, X., Peng, Z., Zhou, Y., Yao, M., Ma, Z., Wang, C., Gao, Z., & Shen, M. (2018). User-defined gestures for gestural interaction: extending from hands to other body parts. *International Journal of Human–Computer Interaction*, *34*(3), 238–250. https://doi.org/10.1080/10447318.2017.1342943

Choi, E., Kwon, S., Lee, D., Lee, H., & Chung, M. K. (2012, September). Can user-derived gesture be considered as the best gesture for a command?: Focusing on the commands for smart home system. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 56, No. 1, pp. 1253-1257). Sage CA: Los Angeles, CA: SAGE Publications.

Commarford, P. M., Lewis, J. R., Smither, J. A., & Gentzler, M. D. (2008). A comparison of broad versus deep auditory menu structures. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *50*(1), 77. https://doi.org/10.1518/001872008X250665

Davis, M. M., Gabbard, J. L., Bowman, D. A., & Gracanin, D. (2016). Depth-based 3D gesture multi-level radial menu for virtual object manipulation. In *Virtual Reality* (pp. 169–170). Greenville, SC, USA: IEEE.

Ferron, M., Mana, N., & Mich, O. (2019). Designing mid-air gesture interaction with mobile devices for older adults. In Sayago, Sergio (eds.), *Perspectives on human-computer interaction research with older people* (pp. 81–100). Springer.

Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*, *47*(6), 381. https://doi.org/10.1037/h0055392

Gebhardt, S., Pick, S., Leithold, F., Hentschel, B., & Kuhlen, T. (2013). Extended pie menus for immersive virtual environments. *IEEE transactions on visualization and computer graphics*, *19*(4), 644–651. https://doi.org/10.1109/TVCG.2013.31

Gran, E. (2018)*Information architecture in vehicle infotainment displays*. [Doctoral dissertation], Arizona State University.

Henderson, J., Malacria, S., Nancel, M., & Lank, E. (2020, April). Investigating the necessity of delay in marking menu invocation. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–13). Honolulu, HI, US.

Henderson, J., Mizobuchi, S., Li, W., & Lank, E. (2019, October). Exploring cross-modal training via touch to learn a mid-air marking menu gesture set. *Proceedings of the 21st International Conference on Human-Computer Interaction with Mobile Devices and Services* (pp. 1–9). Taipei, Taiwan, China.

Hesselmann, T., Flöring, S., & Schmitt, M. (2009, November). Stacked Half-Pie menus: Navigating nested menus on interactive tabletops. *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces* (pp. 173–180). Banff, Alberta, Canada: ACM.

Huckauf, A., & Urbina, M. H. (2008). Gazing with pEYEs: Towards a universal input for various applications. In *Proceedings of the 2008 symposium on Eye tracking research & applications* (pp. 51–54). Savannah, Georgia, USA: ACM.

Institute of Language Teaching, Beijing Language and Culture University. (1986). Modern Chinese frequency dictionary. *Beijing Language and Culture University Press* (pp. 657–658).

Jacko, J. A., & Salvendy, G. (1996). Hierarchical menu design: Breadth, depth, and task complexity. *Perceptual and Motor Skills*, *82*(3_suppl), 1187–1201. https://doi.org/10.2466/pms.1996.82.3c.1187

Kang, B., Choi, E., Kim, H., Lee, M., Oh, S., Jeong, D. Y. et al. (2013). Research on the gesture design tendency of UI experts and users: Focusing on domestic appliances. 대한인간공학회 학술대회논문집, 102-109.

Kiger, J. I. (1984). The depth/breadth trade-off in the design of menu-driven user interfaces. *International Journal of Man-machine Studies*, *20*(2), 201–213. https://doi.org/10.1016/S0020-7373(84)80018-8

Kim, Y., Lee, G. A., Jo, D., Yang, U., Kim, G., & Park, J. (2011). Analysis on virtual interaction-induced fatigue and difficulty in manipulation for interactive 3D gaming console. In *Consumer Electronics (ICCE), 2011 IEEE International Conference on* (pp. 269–270). Las Vegas, NV, USA: IEEE.

Komerska, R., & Ware, C. (2004). A study of haptic linear and pie menus in a 3D fish tank VR environment. In *Haptic Interfaces for Virtual Environment and Teleoperator Systems, 2004. HAPTICS'04. Proceedings. 12th International Symposium on* (pp. 224–231). Chicago, IL, USA: IEEE.

Koutsabasis, P., & Vogiatzidakis, P. (2019). Empirical research in mid-air interaction: A systematic review. *International Journal of Human–Computer Interaction*, *35*(18), 1747–1768. https://doi.org/10.1080/10447318.2019.1572352

Kurtenbach, G., & Buxton, W. (1993). The limits of expert performance using hierarchic marking menus. In *Proceedings of the INTERACT'93 and CHI'93 conference on Human factors in computing systems* (pp. 482–487). Amsterdam, The Netherlands: ACM.

Kurtenbach, G. P. (1993). *The design and evaluation of marking menus*. University of Toronto.

Kurtenbach, G. P., Sellen, A. J., & Buxton, W. A. (1993). An empirical evaluation of some articulatory and cognitive aspects of marking menus. *Human-Computer Interaction*, *8*(1), 1–23. https://doi.org/10.1207/s15327051hci0801_1

Lee, E., MacGregor, & James. (1985). Minimizing user search time in menu retrieval systems. *Human Factors the Journal of the Human Factors & Ergonomics Society*, *27*(2), 157–162. https://doi.org/10.1177/001872088502700203

Lee, M., Kwahk, J., Han, S. H., & Lee, H. (2020). Relative pointing interface: A gesture interaction method based on the ability to divide space. *International Journal of Industrial Ergonomics*, *75*, 102878. https://doi.org/10.1016/j.ergon.2019.102878

Liang, S. F. M. (2013). Control with hand gestures in home environment: A review. *Proceedings of the Institute of Industrial Engineers Asian Conference 2013. Springer Singapore*.

Lim, Z. H., & Kristensson, P. O. (2019, October). An Evaluation of Discrete and Continuous Mid-Air Loop and Marking Menu Selection in Optical See-Through HMDs. *Proceedings of the 21st International Conference on Human-Computer Interaction with Mobile Devices and Services* (pp. 1–10). Taipei, Taiwan, China.

Miller, D. P. (1981). Depth/breadth tradeoff in hierarchical computer menus. *Proceedings of the Human Factors & Ergonomics Society Annual Meeting*, 25, 296–300. NY, USA.

Mo, F., & Zhou, J. (2018). The influence of menu structure and layout on usability of smartwatches. *International Journal of Mobile Human Computer Interaction (IJMHCI)*, *10*(1), 1–22. https://doi.org/10.4018/IJMHCI.2018010101

Norman, D. A. (2010). Natural user interfaces are not natural. *Interactions*, *17*(3), 6–10. https://doi.org/10.1145/1744161.1744163

Norman, K., & Chin, J. (1988). The effect of tree structure on search in a hierarchical menu selection system. *Behaviour & Information Technology*, *7*(1), 51–65. https://doi.org/10.1080/01449298808901862

Paap, K. R., & Roske-Hofstrand, R. J. (1986). The optimal number of menu options per panel. *Human Factors the Journal of the Human Factors & Ergonomics Society*, *28*(4), 377–385. https://doi.org/10.1177/001872088602800401

Pang, X., Guo, R., Yao, N., Yu, J., Wang, C., & Gao, Z. (2014). Human factor studies on gestural interaction: Past, present, and future. *Chinese Journal of Applied Psychology*, *20*(3), 243–251. http://www.appliedpsy.cn/CN/Y2014/V20/I3/243

Pereira, A., Wachs, J. P., Park, K., & Rempel, D. (2015). A user-developed 3- D hand gesture set for human–computer interaction. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *57*(4), 607–621. https://doi.org/10.1177/0018720814559307

Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological review*, *63*(2), 81–97. https://doi.org/10.1037/h0043158

Seppala, P., & Salvendy, G. (1985). Impact of depth of menu hierarchy on performance effectiveness in a supervisory task: Computerised, flexible manufacturing system. *Human Factors*, *27*(6), 713–722. https://doi.org/10.1177/001872088502700610

Shneiderman, B. (2010). *Designing the user interface: Strategies for effective human-computer interaction*. Pearson Education India.

Snowberry, K., Parkinson, S. R., & Sisson, N. (1983). Computer display menus. *Ergonomics*, *26*(7), 699–712. https://doi.org/10.1080/00140138308963390

Tapia, M. A., & Kurtenbach, G. (1995). Some design refinements and principles on the appearance and behavior of marking menus. In *Proceedings of the 8th annual ACM symposium on User interface and software technology* (pp. 189–195). Pittsburgh, PA, USA.

Tian, F., Ao, X., Wang, H., Setlur, V., & Dai, G. (2007, April). The tilt cursor: Enhancing stimulus-response compatibility by providing 3d orientation cue of pen. *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 303–306). San Jose, CA, USA: ACM.

Trinh, H., Waller, A., Vertanen, K., Kristensson, P. O., & Hanson, V. L. (2014, October). Phoneme-based predictive text entry interface. In *Proceedings of the 16th international ACM SIGACCESS conference on Computers & accessibility* (pp. 351–352). Rochester, NY, USA: ACM.

Urbina, M. H., Lorenz, M., & Huckauf, A. (2010). Pies with EYEs: The limits of hierarchical pie menus in gaze control. *Symposium on Eye-Tracking Research & Applications* (pp.93–96). Austin, TX, USA: ACM.

Walter, R., Bailly, G., Valkanova, N., & Müller, J. (2014, September). Cuenesics: Using mid-air gestures to select items on interactive public displays. *Proceedings of the 16th international conference on Human-computer interaction with mobile devices & services* (pp. 299–308). Toronto, ON, Canada.

Wilson, A., & Shafer, S. (2003, April). XWand: UI for intelligent spaces. *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 545–552). Ft. Lauderdale, Florida, USA.

Zhao, S., & Balakrishnan, R. (2004). Simple vs. compound mark hierarchical marking menus. *Proceedings of the 17th annual ACM symposium on User interface software and technology* (pp. 33–42). Santa Fe, New Mexico, USA: ACM.

Zhao, Y., Du, N., Xu, X., Gu, Q., Wang, L., Gao, Z., & Wang, C. (2014). The influence of user-tracking feedback format on gestural interaction's user experience: A Kinect-based usability study. *Chinese Journal of Applied Psychology*, *20*(4), 367–374. http://www.appliedpsy.cn/CN/Y2014/V20/I4/367

About the Authors

Wenmin Li is a graduate student at Zhejiang University, Hangzhou. He majors in applied psychology, obtained Bachelor of Science from Zhejiang University, Hangzhou, China, 2018.

Xueyi Wan is a graduate student at Zhejiang University, Hangzhou. She majors in cognitive psychology, obtained Bachelor of Science from Zhejiang University, Hangzhou, China, 2017.

Yanwei Shi is a graduate student at Zhejiang University, Hangzhou. He majors in applied psychology, obtained Bachelor of Science from Zhejiang University, Hangzhou, China, 2016.

Nailang Yao is a graduate student at Zhejiang University, Hangzhou. He majors in cognitive psychology, obtained Bachelor of Science from Zhejiang University, Hangzhou, China, 2015.

Ci Wang is a lecturer at Zhejiang University, Hangzhou. He majors in engineering psychology, and obtained Master of Science from Zhejiang University, Hangzhou, China, 2011.

Zaifeng Gao is a professor at Zhejiang University, Hangzhou. He majors in cognitive psychology, and obtained Ph.D from Zhejiang University, Hangzhou, China, 2009. His research focuses on engineering psychology and cognitive psychology.