Int. J. Human–Computer Studies 170 (2023) 102952

Contents lists available at ScienceDirect

International Journal of Human - Computer Studies

journal homepage: www.elsevier.com/locate/ijhcs

Effects of spatial constraints and ages on children's upper limb performance in mid-air gesture interaction

Fei Lyu [a,b,*], Huijing Li [a], Qiang Fu [a], Yujie Liu [a], Jin Huang [c], Zhigang Deng [d]

[a] School of Digital Media and Design Arts, Beijing University of Posts and Telecommunications, Beijing, 100876, China
[b] Beijing Key Laboratory of Network System and Network Culture, Beijing University of Posts and Telecommunications, Beijing, 100876, China
[c] State Key Laboratory of Computer Science, Institute of Software, Chinese Academy of Sciences, Beijing, 100190, China
[d] Department of Computer Science, University of Houston, Houston, 77204-3010, USA

ARTICLE INFO

Keywords:
Mid-air gesture interaction
Children
Pie menu
Fitts' law

ABSTRACT

Through two controlled experiments, including a pie menu study and a target acquisition study, this paper investigates children's performance on mid-air gesture interactions in different spatial constraints (i.e., in different orientations/distances), as well as the age effect on such interaction scenarios. The first experiment recorded children's speeds and accuracies following certain directions under menus with different numbers of items, while the second evaluated the speed–accuracy trade-off (SAT) of children's arm movements. We also compared performance differences between two age-related groups (i.e., 6–8 years old and 9–12 years old). Based on these experiments, we propose an improved design for UI menus based on mid-air gesture interaction for children. The improved design provides suggestions for setting appropriate directions and difficulty indexes, which makes it much easier and quicker for children to use the menus with mid-air interaction.

1. Introduction

With the rapid development of low-cost gesture tracking systems, the research of mid-air interaction as a new class of natural user interface (NUI) has become a hot spot in recent years. In some scenarios like large screen interaction and VR interaction, gesture is a natural and intuitive way of interpersonal communication, rich in meaning and convenience (Nacenta et al., 2013; Pereira et al., 2015). Moreover, compared with traditional mouse-keyboard interaction, gesture is considered a more natural and easier interaction technique (Abdul-Rashid et al., 2017). As an innate human skill, gesture requires less cognitive load on the user (Carvalho et al., 2018). Therefore, we believe that mid-air gesture interaction is more appropriate for children as it is more intuitive and natural. Body gesture interaction is generally regarded as the next generation of computer mouse (Norman, 2010; Pang et al., 2014). Recent gesture recognition sensors, such as Microsoft Kinect and Leap Motion, allow interaction with pie menus, particularly in gaming applications such as Far Cry 6 (see Fig. 1-a) and Counter-Strike (see Fig. 1-b). By using body gesture sensing devices, users can use hand movements in different orientations to acquire items in pie menus or any targets in other forms.

A rising research trend on mid-air interaction is to explore its applications for special user groups, such as the elderly (Muangmoon et al., 2016), children (Lyu et al., 2017), or psychiatric patients (Ruiz-Rodriguez et al., 2019). As one of the largest potential user groups of mid-air gesture interaction systems, children are important study subjects for understanding user abilities and user preferences for mid-air interactions. However, unlike the stable performances of adults when using mid-air interaction systems, children of different ages may have different speeds and accuracies in mid-air gesture interaction tasks. To date, there has been a gradual proliferation of somatosensory interaction products for children users, which are widely active in the gaming and education fields. Many somatosensory gaming platforms have gone live with commercial games for children with mid-air gesture interaction. For example, Kinectimals (see Fig. 1-c) allows players to interact with virtual pets through mid-air gestures, and classic games like Cut the Rope (see Fig. 1-d) that were originally designed for touch screens can now be played in leap motion using gestures. In addition to games, there are also other interactive systems for children (Garcia-Sanjuan et al., 2016; Lyu et al., 2017; Ruiz-Rodriguez et al., 2019). But they still lack adequate prior knowledge and experimental evidence on how well children can master this interaction modality. Namely, due to the lack of quantitative studies, parameter settings in most existing applications are ad hoc and empirical. Design choices in one application cannot trivially be extended to other applications, which limits the broad applicability of new mid-air technologies for children.

* Corresponding author at: School of Digital Media and Design Arts, Beijing University of Posts and Telecommunications, Beijing, 100876, China.
E-mail address: feilyu@bupt.edu.cn (F. Lyu).

https://doi.org/10.1016/j.ijhcs.2022.102952
Received 24 January 2022; Received in revised form 11 October 2022; Accepted 16 October 2022
Available online 21 October 2022

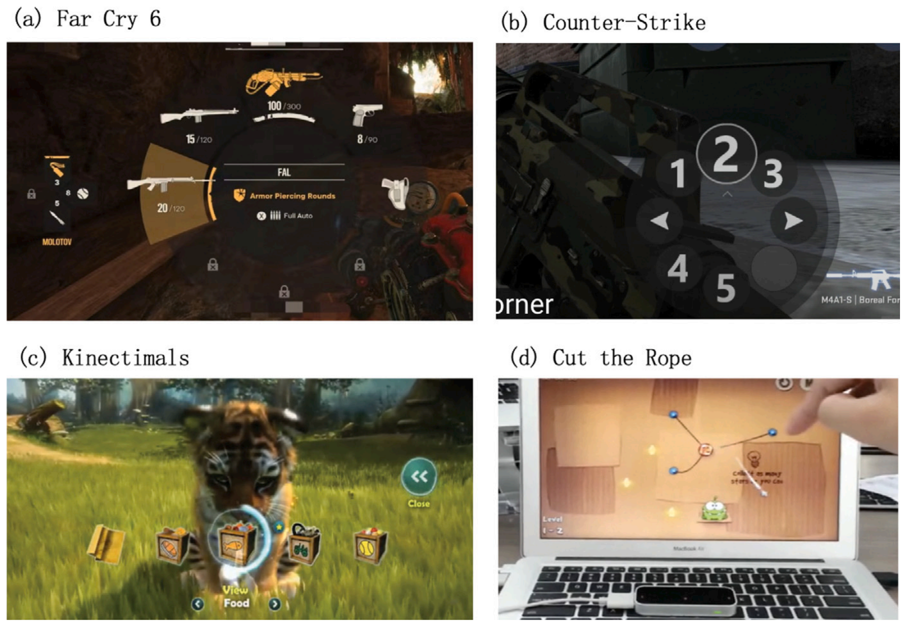

Fig. 1. Several commercial applications that use mid-air gesture interaction: (a) Far Cry 6, (b) Counter-Strike, (c) Kinectimals, and (d) Cut the Rope. The above pictures are from the official websites of these games.

To quantitatively understand children's abilities and preferences on mid-air interaction systems, we conducted two studies focused on analyzing children's arm performances in two widely-used mid-air gesture interaction tasks, i.e., a pie menu task (or pie task) and a target acquisition task. In the pie menu study, we recorded the performance (i.e., speeds and accuracies) of children grouped by age under menus with different numbers of items. We also measured the children's abilities to perform arm movements by kinematic metrics, including variability of movement trajectories and the maximum distance of movement. In the target acquisition study, we further evaluated the speed–accuracy trade-off (SAT) of children's arm movement. This was conducted by comparing the intercepts and slopes of Fitts' law in different directions with children in different age groups. Finally, based on the quantitative and qualitative results obtained from the two studies, we summarize empirical evidence for children's arm movements and derive a set of design guidelines that can be used to design mid-air gesture interaction systems for children.

The remainder of this paper is organized as follows. First, we review the relevant literature on children's mid-air gestural interactions; then, we describe the design, procedures, analyses, and results of two studies on pie menus and target acquisition tasks. Next, we present the results of our studies and summarize the implications for the design of child-oriented mid-air gesture interaction. Finally, we discuss our conclusions and future work.

2. Related work

In this section, we divide the previous work related to mid-air gesture interaction for children into three categories: mid-air upper limb movement, interactive systems for children, and Fitts' law and throughput.

2.1. Mid-air upper limb movement

The human performance of adults, especially adults' upper limb-based motor skills, has been extensively studied. Balakrishnan and MacKenzie (1997) explored the relative bandwidth of users' limb segments, such as fingers, wrists, and forearms, through experiments. Tian et al. (2017) investigated human motor skills to perform discrete menu selection tasks using arm-stretching movements, through two controlled experiments. Wittorf and Jakobsen (2016) focused on free-hand mid-air gestures for wall-display interaction and learned the user-preferred gesture types in this context. Recently, Bachynskyi and Müller (2020) studied the dynamics of aimed mid-air movements. They found that mid-air movements have more complex dynamics than mouse movements. Besides, researchers have also studied the fatigue of upper limb-based movements for adults. Jang et al. (2017) built a biomechanical upper limb model, a three-compartment fatigue model, and maximum shoulder torque estimation to quantify arm fatigue. Cheema et al. (2020) conducted a user modeling experiment using a biomechanical arm simulation model to synthesize mid-air interaction movements and to predict the associated embodied user experience focusing on subjective fatigue.

A few interaction techniques have been proposed based on the study of motor skills of adults. Mine et al. (1997) explored novel body-relative interaction techniques based on the framework of proprioception, which refers to a person's sense of the position and orientation of his or her body and limbs. Ni et al. (2008) proposed rapMenu, a design for freehand menu selection, by using tilt and pinch gestures. Buschek et al. (2018) proposed an interaction technique that combines arm and wrist rotation gestures with simultaneous key to enhance physical keyboard shortcuts. Lyu et al. (2018) explored the combination of 2D directional gesture and 3D depth arm-stretching gesture, and designed a hierarchical marking menu MagicMark, to extend the selection capability of large screen interactions. In addition, Muangmoon et al. (2016) conducted a pilot study on game menu navigation for the elderly using non-tactile gesture interaction, which is derived from the Kinect standard gesture that has been proved to be "very easy to use" and "very fun".

Prior works have ascertained that children differ from adults in the ways they select targets and make gestures. For example, children would miss onscreen targets more often than adults. Cognitive development may vary considerably across age groups of children, particularly in motor skills and executive functions (Chen et al., 2020). Compared with adults, children have significantly weaker motor speed and precision in completing motor tasks. Age also plays an important role in the speed and efficiency of information processing. The accuracy of children's pointing shows significant age differences, i.e., young children have significantly lower pointing ability than older children (Hourcade, 2022). One reason for the development of motor skills in children is the significant increase in the number of myelin sheaths in the brain between the ages of 6 and 8 years (Feldman, 2008). Children become more aware of their hands' positions in space relative to their bodies (i.e., proprioception) and can engage in more fine-grained interactions (Chen et al., 2020). Yan et al. (2000) reported that younger children (mean age 6.4 years) behaved slow, inconsistent, uneven, and nonlinear arm movements compared to older children (mean age 9.2 years) and adults. Connell et al. (2013) conducted an elicitation study on children's defined gestures with Microsoft Kinect, applying a Wizard-of-Oz approach. Vatavu et al. (2015) calculated the objective measures of the consensus between 1312 body gesture preferences of children using a dissimilarity-consensus method. Some researchers also focused on children's selection tasks. Carvalho et al. (2014) evaluated the performance of specific target audiences, including children, adults, and older adults, across different interaction paradigms. Also, they compared the interaction performances of children and adults in target selection experiments using different interaction devices (Carvalho et al., 2015). However, most of the above studies have not investigated the effect of spatial constraints on children's mid-air gestures in different orientations and age groups.

2.2. Interactive systems for children

The book "Child-Computer Interaction" by JP Hourcade mentions that some specific interactions seem to be easier for children when using mid-air gestures. Also, younger children's pointing skills are not as good as older children's, so that younger children require larger targets than older children in order to reach the same level of accuracy (Hourcade, 2022).

Researchers used mid-air gestures to build learning and play systems for children. Lyu et al. (2017) developed EnseWing, an interactive system that can help children with limited music training to experience instrumental ensemble playing, in which children can play music notes by moving a hand horizontally. Adachi et al. (2013) developed a simulation game called "Human SUGOROKU" that consists of a full-body interaction system displaying vegetation succession. It allows immersive participatory learning for elementary school students. Rubegni et al. (2019) studied incarnation-based touchless gestural interfaces, taking a child–display interaction perspective. Moser and Tscheligi (2015) studied 20 children's play experiences of a popular game named "Cut the Rope" with touch and mid-air gestures.

Researchers have also built interactive systems using mid-air gestures to assist children in storytelling. Lu et al. (2011) proposed a digital storytelling system called ShadowStory to support children in performing their stories using handheld sensors.

In recent years, gestural interaction for special children has also attracted attention in the HCI community. Sanchez et al. (2017) designed a video game named BeeSmart that allows children to draw around pictographs and words on a screen. They used the game to help children with Down Syndrome, who have deficits in eye–hand coordination skills. Ruiz-Rodriguez et al. (2019) developed gesture-based video games to help children with autism who have fine-motor coordination problems.

Although various interaction systems for children have been proposed, as described above, their parameter settings are still ad hoc and empirical. Therefore, it is difficult to evaluate their usability with a uniform criterion, due to the lack of a methodological design for children's user interfaces based on upper limb movements.

2.3. Fitts' law and throughput

The speed–accuracy trade-off is one of the most accepted principles for human motion performance in HCI. In general, the speed–accuracy trade-off means that the more accurate the task to be performed, the longer it takes, and vice versa. The speed–accuracy trade-off is derived from Fitts' law (Fitts, 1954; MacKenzie and Isokoski, 2008), which predicts that the movement time (MT) needed to point to a target is logarithmically related to the ratio of the width (W) of the target and the distance (A) to the target.

Fitts' law was originally established in a 1D pointing task with a stylus and later extended in 2D and 3D spaces, with different devices (Accot and Zhai, 2003; Grossman and Balakrishnan, 2004) and tasks (Accot and Zhai, 2002, 1997), and evaluated in different instruction conditions (MacKenzie and Isokoski, 2008; Zhai et al., 2004).

Fitts' law is considered a standardized evaluation tool to investigate the trade-off between speed and accuracy in pointing behavior. There are two common ways to use Fitts' law in evaluation. The first is to compare the intercept and slope in Fitts' law under different conditions. The intercept defines the intersection on the y-axis and is often interpreted as a constant delay of time in a specified apparatus. The slope, on the other hand, describes the acceleration of movement time increases as the index of difficulty increases. Because these two parameters have their physical meanings in movement time prediction, researchers can measure movement times in different conditions and determine how the conditions affect the parameters in Fitts' law relationship (Accot and Zhai, 2002; MacKenzie and Isokoski, 2008). Second, throughput (TP) is calculated as a measure of human performance. The rationale for using TP to measure user performance is that the act of performing a target selection task is akin to transmitting information through the channel of the user, with higher transmission rates indicating more effective user performance. TP has been widely adopted to evaluate the user performances of different specified user interfaces (MacKenzie, 2015).

In this paper, both the parameters of Fitts' law and the metric of throughput are used to investigate the speed–accuracy trade-off of children's arm movements.

3. Pie menu study

The purpose of the first study on a pie menu task is three-fold:

(1) Recording targeting performance and motor characteristics such as speed, accuracy, and maximum range of motion of children's upper arms for different numbers of items divided in the pie menu;
(2) Studying the relation between the performance of children and the directions of upper limb movement;
(3) Analyzing the study results based on ages.

Note that we will further leverage the results of the pie menu study as a prior in the second study (Section 4).

3.1. Experimental setup

3.1.1. Task

The task is a formalization of selecting an item in a menu, called *Pie Menu*. The formalization in this study inherits the basic idea of the pie menu, in which the users wave their hands along a specified direction in the space to finish the task, as illustrated in Figs. 2 and 3. The selectable items are sectors uniformly distributed around the center. For example, a pie menu with 4 items divides the circular space into four equal regions. The orientations of the centerlines of these regions are 0°, 90°, 180°, and 270°, respectively. In our implementation, the boundaries of menu items are hidden to avoid participants using visual feedback to adjust their hand movements. Instead, a thin black arrow in the middle of a menu item is displayed to hint to participants the direction in which they should wave their hands to select the corresponding item. Aiming to study children's arm performances in different directions, we ask the participants to stretch their arms as far as possible in certain directions. Meanwhile, we record the maximum distance each child could reach in each direction, as well as his/her time cost and accuracy.

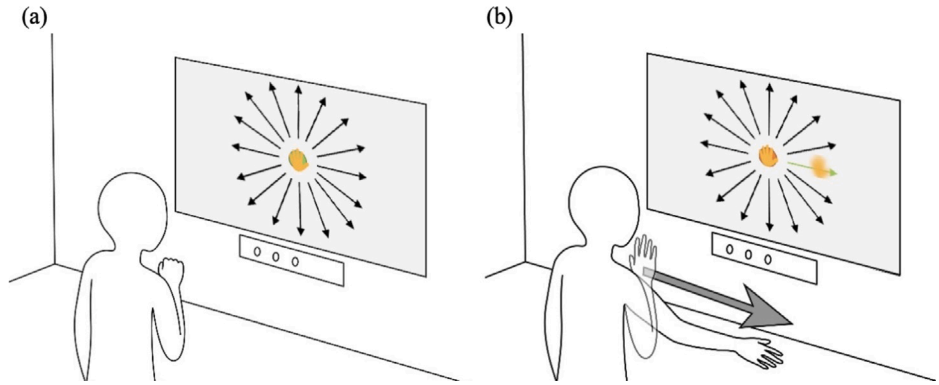

Fig. 2. Illustration of the pie menu experimental task: (a) The user makes a fist for calibration, and (b) the user waves the arm in the specified direction.

Fig. 3. Photos of the operating scenes.

3.1.2. Participants and apparatus

We recruited 20 children (10 boys and 10 girls) between the ages of 6 and 12. To investigate the effect of different age groups on mid-air arm swing performance, we divided the children into two age groups: 10 children aged 6–8 years old (3 children aged 6, 3 children aged 7, and 4 children aged 8) and 10 children aged 9–12 years old (3 children aged 9, 2 children aged 10, 3 children aged 11, 2 children aged 12). We communicated adequately with the child participants and their guardians to ensure that they understood the purpose and content of the study. After obtaining the consent of the participants, they signed an informed consent form.

We conducted the experiment on a Lenovo ThinkPad laptop computer with a 12.5-inch LED display at 1366 × 768 resolution. The used motion-tracking device was a Microsoft Kinect v2.0. All children participating in the experiment were right-handed and did not have prior experience with Kinect or similar devices.

3.1.3. Design

A within-subject factorial design is employed. To explore the capacity limit of children on item numbers in a pie menu, we set the maximum number of menu items to 16. As a result, three menus with 4, 8, and 16 items for different waving directions were used in our experiments. Trials of the three menus are arranged in 3 blocks with 10 repetitions for each direction. The order of the three blocks is counterbalanced with a Latin square design. In each block, the orders of all directions in a menu are randomly provided. Prior to the experiment, each participant has 5 min to practice and there is a ten-minute interval between each block. On average, each participant spent 40 min to complete the whole experiment. In total, we collected data from 5600 trials by 20 participants ((4 direction + 8 direction + 16 direction) × 10 repetitions × 20 participants).

The body shapes of the participants were slightly different, especially their arm lengths. Therefore, before the first waving trial, we asked the participants to calibrate the menu center to the starting position of their dominant hand in a waving gesture. We guided the participants to set the starting position of their waving gesture in front of their shoulders and allowed small deviations to ensure their natural and comfortable feeling. To calibrate the system, the participants raised and held their right hands (Note that all of them were right-handed) in front of their shoulders for 2 s. A “peng” sound was prompted to indicate the calibration was completed (refer to Fig. 2(a)). During the experiment, the participants were asked to keep their bodies as stationary as possible to keep the consistency of their bodies and facing directions.

After calibration, the participants were asked to start the waving test in the direction prompted on the screen. Specifically, the participants placed one of their hands in the center of the starting point for 0.8 s, then the green circle turned red with a “beep” sound emitted. One arrow in a certain direction appeared prompting the waving direction in this trial. The participants were asked to wave their arms as far as possible in the specified direction and to complete the task as quickly and accurately as they could (refer to Fig. 2(b)). The task ended once the participants stayed at the farthest position for 1 s. We also had a verification mechanism to ensure that the arm was straight. Specifically, if the angle between the participant’s forearm flexion and extension was less than 105° (Nam et al., 2019), we treated it as an invalid trial with a warning tone. In this case, the participant was required to redo this trial. We collected the participant’s hand positions with a sampling rate of 100 Hz in the trial. Examples (participants 03, 04, and 09) of movement trajectories are shown in Fig. 4.

3.1.4. Measurements

Our measure is inspired by the path-based pointing measure of MacKenzie et al. (2001) and Vatavu et al. (2013). MacKenzie et al. defined the task axis as a straight line between the user’s mouse starting point and the pointing target, against which the accuracy of the user’s pointing path is compared (MacKenzie et al., 2001). Without loss of generality, we call the vector from the starting position to the end position of the participant’s waving trajectory as the *waving vector*, and the vector of each given direction in the task as the *direction vector*.

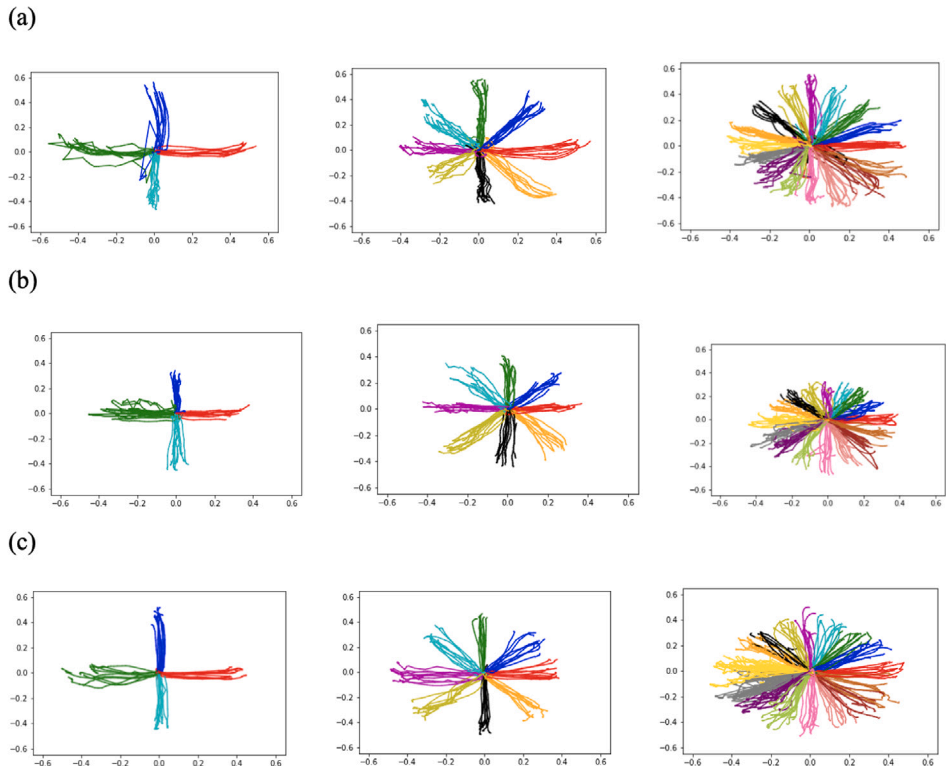

Fig. 4. Example motion trajectory diagrams: (a) (b) (c) represent the participants 03, 04 and 09, respectively. The first to third columns (from left to right) for each participant represent the trajectory diagrams for the 4-, 8- and 16-item menus, respectively.

To quantitatively analyze the experimental results, we also define the following measures:

- *Completion time*: the time required to complete a trial, from the confirmation of the initial gesture to the end of the waving.
- *Offset*: The average offset of the trajectory points of a waving movement perpendicular to the direction vector given in the task.
- *Range of motion*: the maximum waving distance of one trail.
- *Error rate*: An error is defined if the angle between the waving vector and the direction vector in the task is larger than a specified threshold. The error rate is calculated by dividing the number of errors by the total number of trials in a condition. The thresholds of the three menus with 4, 8, and 16 items, are set to ±45°, ±22.5°, and ±11.25°, respectively.

Movement time (MT) and error rate (ER) are two common measures of pointing speed and accuracy. However, the two measures are not good at capturing subjects' motor behavior during the experiment (MacKenzie et al., 2001). Therefore, we introduce the motion offset metric (Vatavu et al., 2013) to reveal the characteristics of the gesture motion path.

3.2. Experimental results

We analyze the experimental results in three aspects: (1) The quantitative performance of the participants in the pie menu study in terms of completion time, offset, range of motion, and error rate. (2) The effect of ages on completion time and error rate was analyzed by comparing the 6–8 years age group with the 9–12 years age group. We also compared the performance of children and the previously reported performance of adults (Ni et al., 2011; Gang and O'Neill, 2012). (3) The effect of different menus on subjective user experience. We run a Kolmogorov–Smirnov normality test at the significance level of 0.05 before statistical analysis. Repeated-measures ANOVAs are used if the data followed a Normal distribution, while Friedman tests are adopted if the data did not follow a Normal distribution. Post hoc multiple analyses are also performed, and Bonferroni correction is used to adjust for the p-value to compare the differences between the two groups.

3.2.1. Quantitative performance

In Table 1, we summarize the pie menu study results in terms of four types of quantitative measures. Note that we focus on the top-2 ("maximum" and "secondary maximum") and last-2 values ("minimum" and "secondary minimum") for each measure to make these values comparable. For each value, we also show its corresponding angle to highlight the impact of angles. The overall average error rate of all participants is 3.81%. Friedman's test shows that there is a significant main effect for Menu, $X^2(2, N = 20) = 32.771$, $p < .001$, and the main effect of Age is marginally significant, $X^2(1, N = 30) = 3.857$, $p = .050$.

As can be seen from Table 1, when the number of items in a pie menu increases, the quantitative performance decrease, i.e., more completion time required and a larger error rate. Moreover, the user performances on pie menus with different items are consistent with corresponding angles. As the number of items increases, the influence of the angles is more accurate, e.g., the slowest angles are gradually changed from 180° on the 4-item menu to 112.5° on the 16-item menu. Similar scenarios can also be observed in the other three measures. Post hoc analyses show that, for the 4-item menu, there are significant differences between all pairs except for the pair between 90° (43.70 cm) and 270° (44.05 cm) directions ($p = 0.871$). For the 8-item menu, the range of motion in the 0° direction (40.01 cm) is significantly shorter than the other directions (all $p < .01$). For the 16-item menu, pairwise comparison results show that there is no significant difference between any two directions (all $p > .05$). In order to show the influence of directions more clearly, we plot the range of motion of each direction in the 4-, 8- and 16-item menu on a radar map, as shown in Fig. 5.

On the other hand, we can see that the error rate increases sharply from the 8-item pie menu to the 16-item one. It is mainly because the over-divided pie menu makes the items difficult to correctly select.

Table 1
The average completion time, range of motion, offset, and error rate for three different pie menus. In each row, we show the top-two ("maximum" and "secondary maximum") and last-two ("minimum" and "secondary minimum") values for each measure as well as their corresponding angles in the parentheses.

	Maximum	Secondary maximum	Secondary minimum	Minimum
Completion time				
4-items	1519 ms (180°)	1399 ms (90°)	1397 ms (270°)	1370 ms (0°)
8-items	1605 ms (135°)	1572 ms (180°)	1437 ms (315°)	1397 ms (0°)
16-items	1744 ms (112.5°)	1740 ms (135°)	1538 ms (337.5°)	1460 ms (0°)
Range of motion				
4-items	44.05 cm (270°)	43.7 cm (90°)	41.9 cm (180°)	39.98 cm (0°)
8-items	44.48 cm (90°)	43.7 cm (270°)	41.4 cm (45°)	40.01 cm (0°)
16-items	44.12 cm (270°)	43.3 cm (247.5°)	38.2 cm (45°)	38.13 cm (67.5°)
Offset				
4-items	4.01 cm (180°)	3.42 cm (0°)	2.46 cm (270°)	2.21 cm (90°)
8-items	4.03 cm (180°)	3.97 cm (45°)	2.47 cm (270°)	2.42 cm (90°)
16-items	5.94 cm (22.5°)	5.18 cm (67.5°)	2.44 cm (270°)	2.44 cm (90°)
Error rate				
4-items	1.5% (90°)	0.5% (0°)	0.5% (180°)	0% (270°)
8-items	1.5% (90°)	1.5% (225°)	0% (180°)	0% (270°)
16-items	20.5% (22.5°)	15.5% (315°)	3% (270°)	1.5% (90°)

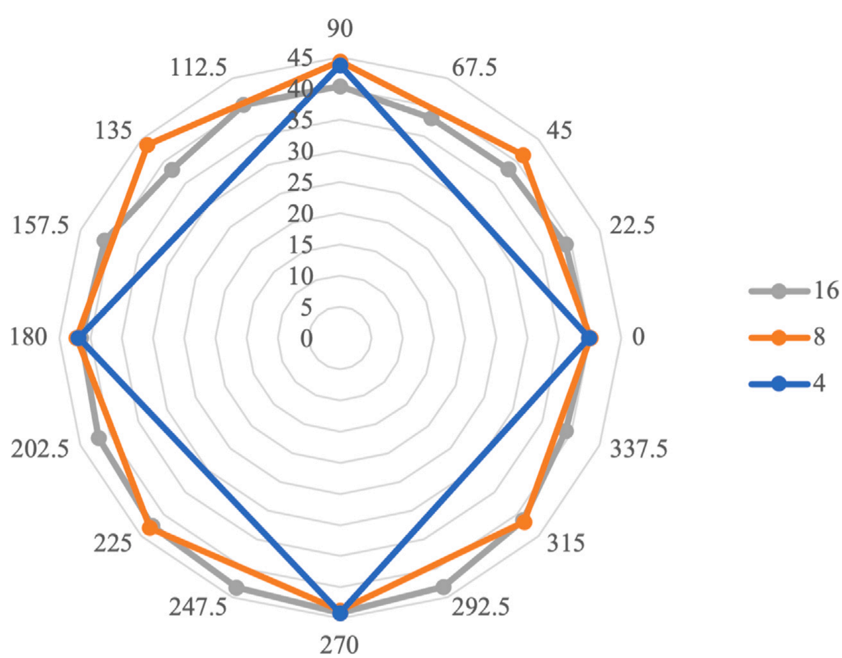

Fig. 5. Maximum range of motion for each direction in the three types of pie menus.

This demonstrates that a user interface involved with a pie menu task should balance functionality and usability. Increasing the number of items could increase the functionality of the interface, but it also could increase the difficulty and error-prone aspect of accessing those functions for children. Our results indicate that, in a full 360° pie menu, the threshold of this balance is the 8-item menu.

3.2.2. Impacts of age

To analyze the impact of the age factor on user performance, we separate the data of the participants into two age-based groups, i.e., groups for 6–8 years old and 9–12 years old, respectively. Fig. 6 shows the means and standard deviations of the error rates in different age groups on the three menus. As for the age factor, post hoc analyses show that the error rates of the 6–8 age group (M = 12.43%) are significantly higher than those of the 9–12 age group (M = 7.31%) for the 16-item menu, $X^2(1, N = 10) = 10.000$, $p = .002$. The error rates between the two age groups for the 4- and 8-item menus are not significantly different ($p > .05$).

We found that in the pie menus with 4 or 8 items, the error rates of the participants are reasonably low. For the two menus, the error rates, completion time, and average offsets of the two age groups do not have statistically significant differences. When the number of items is increased to 16, the error rates of the participants increased significantly, and the two age groups performed quite differently. The main reason could be that, in the condition of a higher item density, both the radial distance and the fault tolerance threshold are lower, which causes the participants to tend to make more errors during the waving tasks, especially for younger children with less developed arm motor skills. The participants' subjective evaluations also confirm the above results.

In addition, we further compare the user performances between children and adults on direction-related, mid-air gesture interaction. As reported in previous research studies on marked menus (Ni et al., 2011; Gang and O'Neill, 2012), many marked menus for adults generally do not have more than 8 items, which is consistent with our findings.

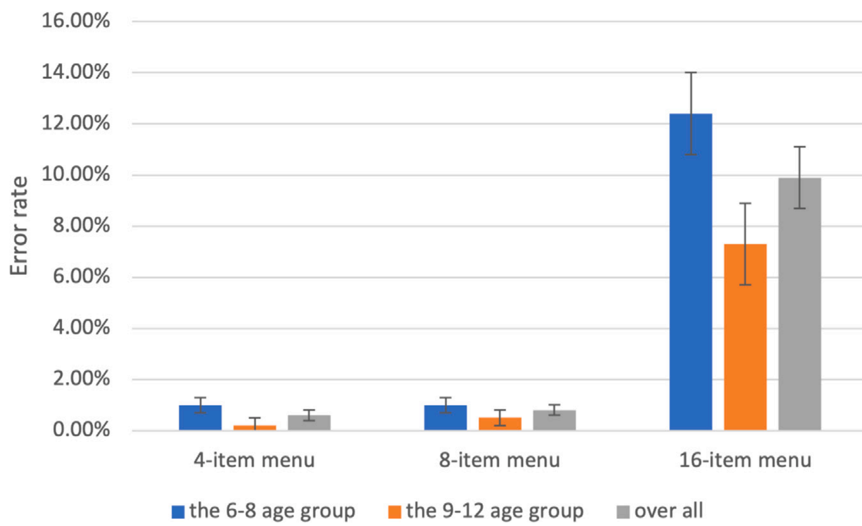

Fig. 6. The mean values and standard errors of the error rates of different age groups for the three types of pie menus.

For example, Ni et al.'s study on rapMenu (Ni et al., 2011) shows that in a circular menu based on wrist-tilt interaction, users select menus with 8 items with the fastest speed and selected menus with 20 items with the slowest speed; when the number of items in a menu is increased from 8 to 16, the error rate is significantly increased (i.e., from below 2% to about 4%). Still, the error rate increase for adults is much smaller than that of children for the same item density cases (from 0.75% to 9.87% for children). Probably, the main reason is that adults have a higher arm movement precision than children in general, so they can still maintain a relatively stable and low error rate when handling more levels of menu items. Gang and O'Neill (2012)'s study of 3D marker menus shows a significant effect of target orientation on selection time and error rate: for adult users, the right orientation is very comfortable, while the up and down orientations are also preferred. This is largely consistent with the experimental results of child users, i.e., users perform well in all three directions, 0°, 90°, and 270°. Their study also reported a subjective dislike of down-left and back-left orientations among adult users. Users had difficulty in selecting targets in the vertical direction, especially downward targets (down, left down, and right down), and there were high error rates, which could be caused by the motion pattern of the arm: hands usually do not move in the same vertical plane as the user's chest (Gang and O'Neill, 2012). Our experimental results show that in an 8-item menu task, child users also have higher error rates in the 90° and 225° directions.

3.2.3. Performances in different directions

To understand children's performances in different directions, we mainly refer to the user performances in the 16-item menu task, as it can better reflect the limit of user performances. As shown in Fig. 7, the radar map of the error rates of the 16-item menu task presents a "bat-like" distribution. That is, in the distribution, the error rates in the up (90°) and down (270°) directions are significantly lower than in the other directions. For the "bat wings" on both sides, the error rates in the middle of the wings (0° and 180°) are relatively low, while the error rates in the oblique directions, such as 45° and 22.5°, are higher. It is interesting to note that, although all of the children who participated in our study are right-handed, between the two "bat wings", the wing on the left side has slightly lower error rates than the wing on the right side. The main reason could be that the children in the 6–8 age group reduced their error rates with the cost of increasing the completion time when completing tasks on the left side, especially on the upper left. As such, the results show lower error rates on the left wing.

Our results also show that, in terms of the error rate, children perform best in the up-and-down direction, followed by the directions close to the top and bottom (i.e., 67.5°, 112.5°, and 292.5°), while in most directions on the right side, children perform worse. Combined with the results of the offset and completion time analysis, we found that children show similar characteristics in controlling their arm movement: accurately controlling the movement when waving up and down, being slow and cautious when waving in the direction on the left side, being casual and fast when completing movement on the right side, which leads to slightly worse performance.

3.2.4. Direction of calibration

For the 8-item menu task, our results show that the range of motion of the participants in the right (0°) direction is the shortest. By contrast, the ranges of motion on the left side (i.e., upper-left, left, and lower-left) are significantly larger than those on the right side (i.e., upper-right, right, and lower-right). This is different from our speculation before the study. We found an interesting issue based on the arm trajectories and our observations during the experiment. Although the participants were explicitly asked not to turn their shoulders and upper bodies during the experiment, more than half of them involuntarily twisted their shoulders to the left, in order to satisfy the requirement of "wave their arms as far as possible" to make their own performance better. Such performances may lead to an increase in the waving distances and completion time in the directions on the left side.

Since the postures of waving to the left cannot be unified into a comfortable posture suitable for all the participants to wave, it is not a suitable direction of mapping. In summary, the right direction can undoubtedly be used as the direction in which the user straightens the arm when mapping the length of the arm in the follow-up "target acquisition study" experiment.

4. Target acquisition study

In the second study, we investigate the speed–accuracy trade-off of children's arm movements in terms of Fitts' law. This goal is achieved through a target acquisition task, in which we focus on the following three key questions:

(1) How do movement orientation and age affect children's motor performance in target acquisition tasks?
(2) Do the performances of target acquisition in children's aim movements match Fitts' law?
(3) If the answer to the above second question is "yes", how does the index of difficulty (ID for short) in Fitts' law vary across different spatial orientations and ages in children's aim movements?

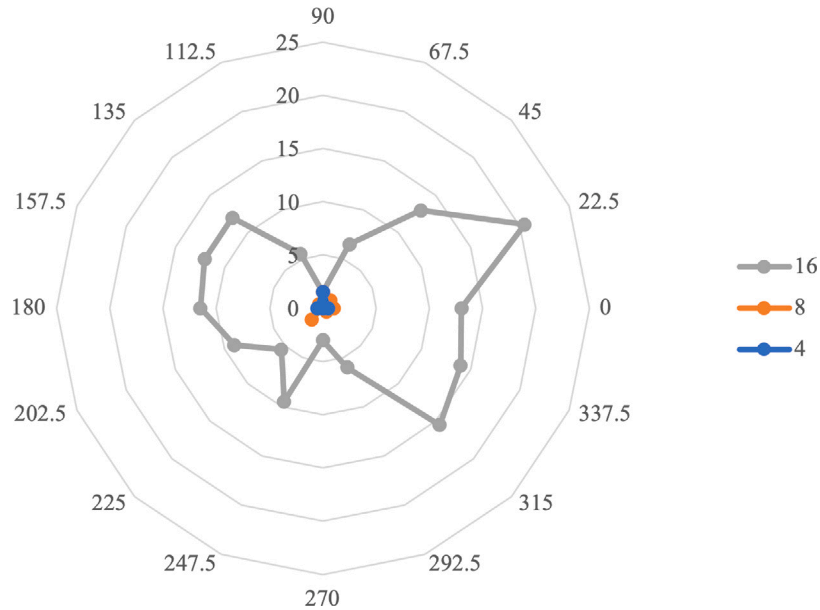

Fig. 7. Error rate for each direction in the three faction menus.

4.1. Experimental setup

4.1.1. Participants and apparatus

We recruited fourteen volunteers and divided them into two groups: (i) seven children were younger (6 to 8 years old), (ii) the other seven children were older (9 to 12 years old). The apparatus used in this experiment is the same as in the first study. All participants are right-handed and have no prior experience in Kinect or similar visual-based interaction devices.

4.1.2. Design

Participants are asked to move one of their hands to control a spot moving from a starting point to a circular target on the screen and hover for a while to select the target. The experiment is carried out as a within-subject design across 72 conditions generated by 3 factors, including *Direction*, *Width*, and *Distance*. The target is set in one of eight different directions, with the width of 50 pixels (1 cm), 75 pixels (1.5 centimeters), or 100 pixels (2 centimeters). The distance from the target to the starting point is set to one of three values, 80 pixels (1.6 centimeters), 160 pixels (3.3 centimeters), or 320 pixels (6.5 centimeters). A total of 72 combinations (i.e., 8 directions × 3 widths × 3 distances) are randomized to reduce the order effect. Before the experiment, each participant has a 5-minute practice session. Participants can take a 10-minute break after 2 repetitions. The experiment is anticipated to last approximately 20 min for each participant on average. In total, we collected data from 4032 trails by 14 participants (8 directions × 3 distances × 3 widths × 4 repetitions × 14 participants).

4.1.3. Procedure

First of all, we normalized the arm length of the participant by asking her/him to raise the right arm horizontally and then mapped her/his arm length to 320 pixels (the longest distance of the target) on the screen. To ensure all children with different arm lengths can reach the target with the maximum distance on the screen, we set such a ratio as $M/D = L/D_{max}$, where M represents the movement in physical space, D represents the distance moved in the screen, L represents the arm length of the child, and D_{max} represents the maximum distance on the screen. For example, if a child's arm length is 60 centimeters, he must move his hand 60 centimeters to reach a target with a maximum distance of 320 pixels, whereas he only needs to move 15 centimeters to reach a target with a target distance of 80 pixels. After the normalization, we asked the participant to raise the right fist in front of the right shoulder for calibration (refer to Fig. 8(a)), which was similar to the calibration process in the first study. Before each experiment, a green circle with a radius of 50 pixels was placed in the center of the screen as the starting point. The child participant's hand was mapped on the screen as a cursor, and we asked the child to hold the cursor on the starting point for about 0.8 s. The green starting point in the center of the menu turned red and a "beep" sounded to indicate the start of an experiment. Meanwhile, a new green circle appeared on the screen at a specific location to represent the target. The child participant needed to move his/her hand to the target as quickly and accurately as possible and stay for about 0.8 s to complete the selection task (refer to Fig. 8(b)). And Fig. 9 shows a photo of the operating scene.

If the participant correctly captures the target, the system will record the trial as a success. Then, the target disappears with a "ding" sound, and the central circle returns to green for the next round. In any of the following cases, the system recorded an error and played an error sound: (i) the participant did not move his/her hand into the target area within 0.8 s after the start of the experiment; (ii) the participant moved his/her hand into the target area but did not stay there for more than 0.8 s. The next round would not start until the participant successfully completes this trial.

After all trials, participants were asked to complete a Likert scale questionnaire that provides subjective ratings of different aspects of each type of goal. Likert scales are widely recognized as one of the simplest and most reliable techniques used for attitude measurement which is appropriate for children (Royeen, 1985). We used a 5-point Likert scale ranging from 1 (worst) to 5 (best). According to Adelson and Mccoach (2010), the 5-point scale is more appropriate for children at the elementary school level. The designed questions include: how easy is it to capture the target points? How fast is it to capture the target? How accurate is it to capture the target? etc. And the questions were all set as straightforward, simple sentences, and the research assistant aided in describing the explanation of the questions to ensure that the participants understood the questions.

4.1.4. Measurements

The following metrics are employed to measure the task performance of the participants:

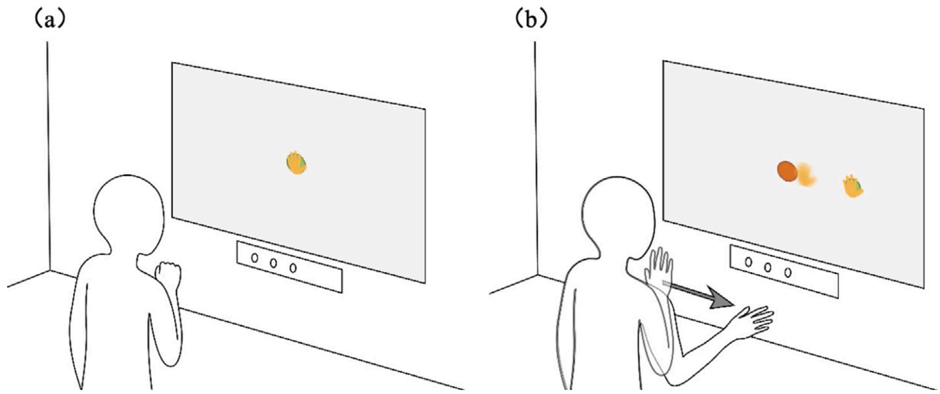

Fig. 8. The target acquisition experimental task: (a) The user makes a fist for calibration; (b) the user moves one hand to capture the green target according to on-screen prompts.

Fig. 9. Photos of the operating scenes.

- *Task completion time*: The time from the appearance of a target to the capture of the target;
- *Error rate*: The percentage of times participants failed to capture the target in all experiments. The errors are considered to be the scenarios when the child participant's hand fails to stay within the target area for 0.8 s, and when the distance from the location where the child participant's hand stays for 0.8 s to the center of the target circle is greater than the radius of the target.

4.2. Experimental results

We report the experimental results in two parts: (1) The effects of the ID factor and the age factor on the completion time and the error rate, respectively, and (2) the fitting of the Fitts' law.

The speed–accuracy trade-off originated from the widely-known Fitts' law (Fitts, 1954; MacKenzie and Isokoski, 2008), which predicts that the movement time (MT) needed to point to a target is logarithmically related to the ratio of the width (W) of the target and the distance (A) to the target:

$$MT = a + b\log_2(A/W + 1), \quad (1)$$

where a and b are regression coefficients (called intercept and slope, respectively). The term $log_2(A/W + 1)$ in Eq. (1) is called the index of difficulty (ID) of the task (MacKenzie and Buxton, 1992), describing the difficulty to complete the task with a specified apparatus (e.g., using a mouse to point at a target displayed on a monitor).

The throughput (TP) is calculated by dividing the index of difficulty by movement time as follows:

$$TP = ID/MT = log_2(A/W + 1)/MT \quad (2)$$

4.2.1. Completion time

In Fig. 11, we summarize the completion time for different IDs. The average completion time of all the child participants is 1680 ms. A repeated-measures ANOVA shows significant effects of ID ($F_{6,72}$ = 215.800, $p < .001$), Distance ($F_{2,24}$ = 387.140, $p < .001$), Width ($F_{2,24}$ = 61.319, $p < .001$) on completion time. Post hoc analysis shows that the completion time is significantly different between all pairs of IDs ($p < .001$) except the one between ID = 2.070 and ID = 2.397 ($p = .135$). The task that took the longest time is ID = 2.888 (2300 ms), while ID = 0.848 took the shortest time (1074 ms).

We further look into the data by comparing the completion time of different age groups in separated ID conditions using a paired sample t-test. As shown in Fig. 10, in the case of the lowest index of difficulty (ID = 0.848), the completion time for children aged 6–8 is 1141 ms, which is higher than the 1008 ms completion time for children aged 9–12, $F_{1,12}$ = 5.033, $p = .049$. However, the differences in completion time between the two age-related groups are small. The results of both groups show the consistency of their performances on IDs, that is, the

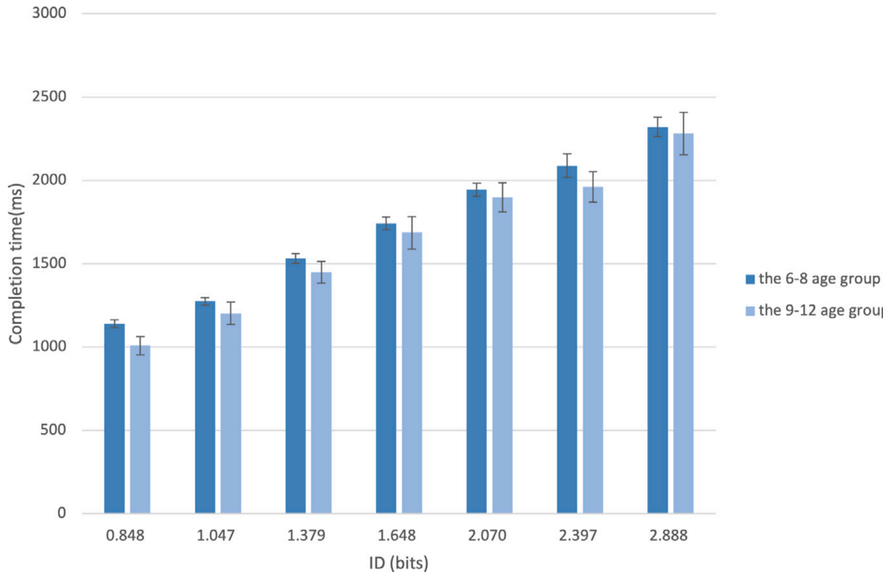

Fig. 10. Completion times of different age groups in separated ID conditions. Error bars indicate the standard errors.

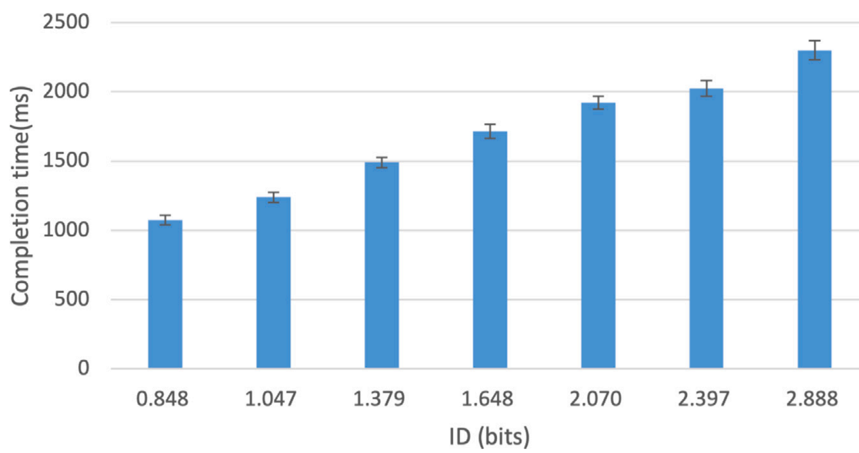

Fig. 11. Completion times for all child participants in the target acquisition study, with respect to different IDs. Error bars indicate the standard errors.

higher the index of difficulty, the more completion time required. From Fig. 11, we can see an approximately linear increasing trend of the completion time with the increase of the ID.

4.2.2. Error rate

Analogously, we illustrate the pair-wise error rates for different IDs in Fig. 12. The overall average error rate of all participating children is 4.17%. Friedman tests show significant effects of ID (X^2(6, N = 14) = 57.312, $p < .001$), Distance (X^2(2, N = 14) = 20.415, $p < .001$), and Width (X^2(2, N = 14) = 24.275, $p < .001$) on the error rate. Post hoc analysis shows that the error rates are significantly different between the condition of the highest index of difficulty (ID = 2.888, error rate = 16.07%) and the others (ID = 0.848, ID = 1.047, ID = 1.379, and ID = 1.648). Different from the completion time that has an approximately linear relation to the ID (refer to Fig. 11), we can see that there exists a sharp increase in the error rate from ID = 1.648 to ID = 2.070. Hence, the error rate has a non-linear relation to the ID.

We further look into the data by comparing the error rates of different age groups in separated ID conditions (Fig. 13) using Friedman tests. In the condition of the highest index of difficulty (ID = 2.888), the average error rate for the 6–8 age group is 20.54%, which is much higher than the 11.61% for the 9–12 age group, X^2(1, N = 7) = 3.571, $p = .059$. However, no statistically significant main effect of the age factor is found, X^2(1, N = 49) = .125, $p = .724$. However, one interesting observation is that the relative performances of two age-related groups are reversed after the step point. Specifically, when ID $\leq$ 1.648, the error rate of the 6–8 age group is lower than that of the 9–12 age group. However, when ID $\geq$ 2.070, the 9–12 age group performs better than the 6–8 age group. This could be explained as follows: The step change of the error-rate vs. ID relation impacts the younger participants more.

4.2.3. The Fitts' law fitting results

Fig. 14 shows the fitting results of Fitts' law for the two age groups. We observe high fits of Fitts' law model with R^2 values of 0.9852 and 0.9730 for the 6–8 age group and the 9–12 age group, respectively. The slopes of the models for the two age groups do not show much difference, but the intercept for the 6–8 age group is slightly higher than the 9–12 age group. The throughput of the 6–8 age group for each direction is lower than the 9–12 age group. In the 6–8 age group, the lowest throughput is 7.45 bits/s for ID = 0.848 and the highest is 12.54 bits/s for ID = 2.888. In the 9–12 age group, the lowest throughput is 8.40 bits/s for ID = 0.848 and the highest is 12.83 bits/s for ID = 2.888.

In order to verify the fittings of Fitts' law, we mainly focus on the conditions of different directions and different age groups. It is obvious that linear regression achieves high fitting accuracy for data divided

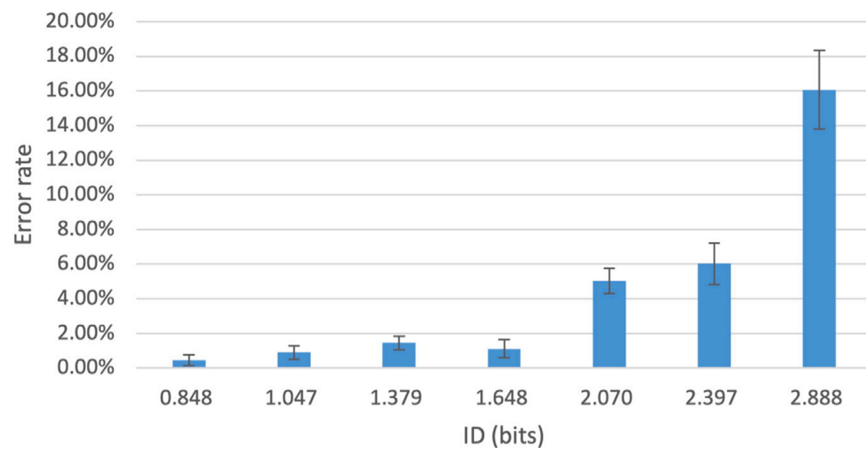

Fig. 12. Error rates for all child participants in the target acquisition study with respect to different IDs. Error bars indicate the standard errors.

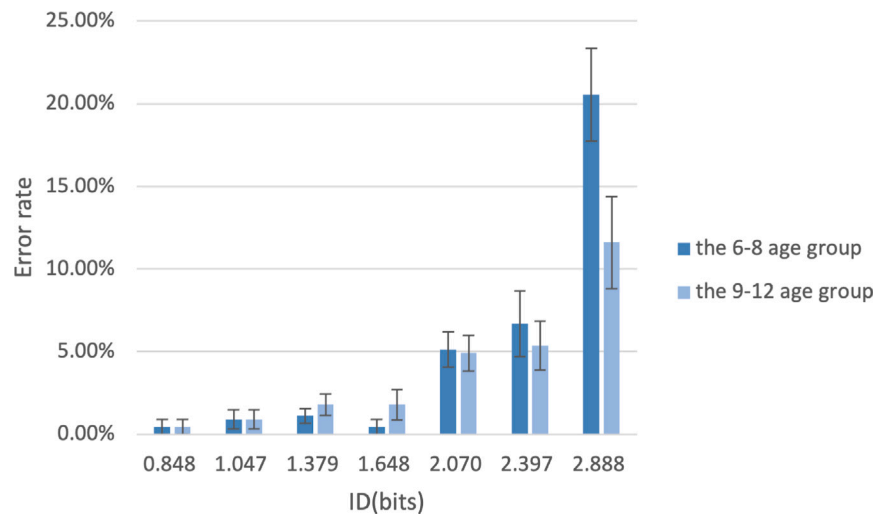

Fig. 13. Error rates of different age groups in separated ID conditions. Error bars indicate the standard errors.

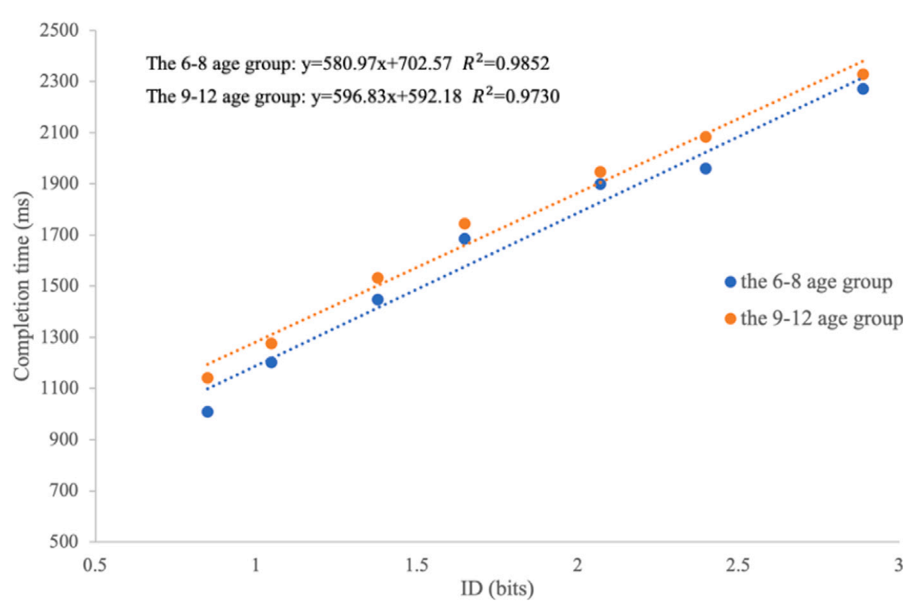

Fig. 14. The fitting results of Fitts' law for the two age groups.

Table 2
Table of the throughputs of the mid-air target selection tasks for the eight directions.

DIR	ID							
	0.848	1.047	1.379	1.648	2.070	2.397	2.888	AVG
0°	8.57	8.72	9.02	9.12	11.08	12.05	12.97	10.21
45°	7.72	8.48	9.28	9.46	11.30	12.17	12.78	10.17
90°	7.96	8.31	9.54	9.83	11.29	11.57	13.59	10.29
135°	7.48	8.10	8.92	9.36	10.19	11.55	12.62	9.74
180°	7.20	8.45	9.25	9.87	10.68	10.89	10.58	9.56
225°	8.35	8.35	9.25	9.69	9.94	11.33	12.54	9.92
270°	8.02	8.61	9.37	10.39	10.87	12.68	14.26	10.60
315°	7.93	8.43	9.30	9.49	11.16	12.2	12.85	10.19
AVG	7.90	8.43	9.24	9.65	10.81	11.80	12.77	

according to 8 directions and 2 age groups with no fewer than 93% of the variances explained by the model. This indicates that the target acquisition task for children's mid-air gestures still conforms to the law of speed and accuracy (i.e., Fitts' law). That is, when the target is smaller and the distance is farther, it takes longer to complete the task, and vice versa.

Moreover, by employing Eq. (2), we calculate the throughputs of the mid-air target selection tasks for the eight directions. As we can see from Table 2, increasing ID increases the throughputs of mid-air target selection. The average throughput in the condition of ID = 2.888 is the highest (12.77 bit/s) while the average throughput in the condition of ID = 0.848 is the lowest (7.90 bit/s). Among all the directions, the highest throughput is found in the 270° direction with an average value of 10.60 bit/s, while the lowest is found in the 180° direction with 9.56 bit/s. As a result, we observe that under all conditions, the throughput of ID = 2.888 in the 270° direction is the highest (14.26 bit/s), and that of ID = 0.848 in the 180° direction is the lowest (7.20 bit/s). Friedman tests show significant effects of ID on subjective ranking of the error-prone feature ($X^2(6, \mathrm{N} = 14) = 53.157$, $p < .01$), speed ($X^2(6, \mathrm{N} = 14) = 63.834$, $p < .01$), and ease of use ($X^2(6, \mathrm{N} = 14) = 63.515$, $p < .01$).

4.3. Discussion

As shown in Fig. 15, the subjective rankings of the error-prone feature range from 1.71 (ID = 0.848) to 4.57 (ID = 2.888). The subjective rankings of the speed range from 2 (ID = 2.888) to 4.64 (ID = 0.848). The subjective rankings of the ease of use range from 1.64 (ID = 2.888) to 4.85 (ID = 0.848). Post hoc analysis shows that most of the pairwise comparisons are significantly different ($p < .01$). The results of subjective rankings for different IDs are consistent with the evidence of task completion time and error rate. The ranking of the error-prone feature increases as the index of difficulty increases. On the other hand, the ranking of the speed factor decreases as the index of difficulty increases.

On the other hand, since the definition of the index of difficulty (ID) involves only two factors (i.e., distance and target width), we also investigate the participants' performances in different directions. As shown in Fig. 11, the completion time of different IDs ranges from 1074 ms (ID = 0.848) to 2300 ms (ID = 2.888). As shown in Fig. 16, for the directions, the completion time ranges from 1602 ms (270°) to 1748 ms (180°), with no main significant effect found, $F_{7,84} = 1.399$, $p > .05$. The 180° direction (1748 ms) is significantly higher than the 270° direction (1602 ms) ($p = .034$). We also illustrate the impact of the directions on the error rate in Fig. 17. The error rates range from 2.78% (45°) to 5.56% (180°), with no main significant effect found, $X^2(7, \mathrm{N} = 14) = 7.126$, $p = .416$.

Furthermore, there are several important values in fitting Fitts' law, including intercept, slope, and throughput. For different directions, the largest slope occurs in the 180° direction, while the smallest one occurs in the 90° direction. This indicates that, in the 180° direction, the time needed for children to perform mid-air selection could increase quickly with the increase of ID. While in the 90° direction, the time needed for children to perform mid-air selection increases relatively slowly with the increase of ID. Interestingly, although the average completion time in the 180° direction is obviously larger, it has the smallest value of intercept. This result could be due to two main reasons. First, it is compensation for the high slope estimated by the linear regression. Second, it could be also caused by the natural starting posture of the participants. We observe that the participants tend to have their right hands heading to left before starting the trial, which leads to a faster reaction time to the target on the left side than that on the right side. For the throughput, it is obvious that in different directions, the TP in the 180° direction is generally smaller than those in the other directions, and the TP in the 270° direction is generally higher than those in the other directions. The above results suggest that children's index of performance in the 180° direction is the worst and the best in the 270° direction.

For the two studied age groups (6–8 ages, and 9–12 ages), we did not find a statistically significant difference between slopes. In terms of the intercept, the 6–8 age group has a larger intercept than the 9–12 age group. This indicates that it might take more reaction time for younger children to perform this task. In sum, in terms of the throughput, the 9–12 age group performs obviously better than the 6–8 age group.

5. Design implications

Based on the empirical evidence in this study, we derive the following user interface design implications for children involving mid-air gesture interaction:

1. The number of mid-air menu items had the greatest impact on children's performance. Children in both age groups had low error rates (≤1%) when using 4-item and 8-item menus. In contrast, children took longer time and had higher error rates on the 16-item menu, but the older children had significantly lower selection error rates than the younger children. In the real-world design environment, user interfaces involving pie menu tasks must strike a balance between functionality and usability. Increasing the number of pie menu items would enhance the functionality of the interface, but would also make it more difficult and error-prone for child users to access these items. In particular, children have a significantly higher error rate in 16-item menus than adults. An informal analysis of the menu widths of popular applications by Bailly et al. (2008) showed that a quarter of them had more than 16 items. Perhaps in order to accommodate too many menu functions, adults may choose to use 16-item menus without creating a poor experience, but it is not appropriate to use 16-item menus in a child-oriented pie menu design. For children, the threshold for functionality and usability balance in a full 360° pie menu is 8 menus. If a pie menu must contain more than 8 items, designers should give special consideration to the reduced accuracy of selection for younger children.

2. In general, children performed the pie menu task more accurately in the upward (90°) and downward (270°) directions, followed by directions closer to the north and south (67.5°, 112.5°, 292.5°, etc.). The results of fitting Fitts' law model showed that children had the worst performance index in the 180° direction. It can be seen that children show the same characteristics in controlling arm movements: accurate control of movement during up and down swings, slow and cautious swings in the left direction, and tend to be random and fast for swings in the right direction. Based on these observations, designers should place objects that require higher operational precision in the up and down directions. Some convenient operations that need to be performed quickly or objects that need to be accessed as soon as possible can be placed on the front right side, while the front left side is suitable for some operations that are not frequently used or need to be performed carefully. However, it should be noted that this conclusion may be affected by the fact that all experimental participants are right-handed, and designers should take it into account in the actual design work.

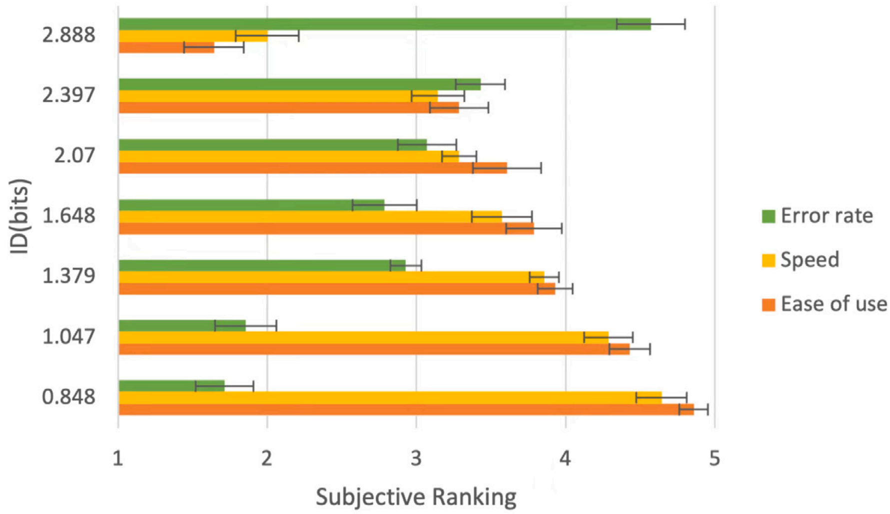

Fig. 15. Subjective rankings of all the participants for different IDs. Error bars indicate the standard errors.

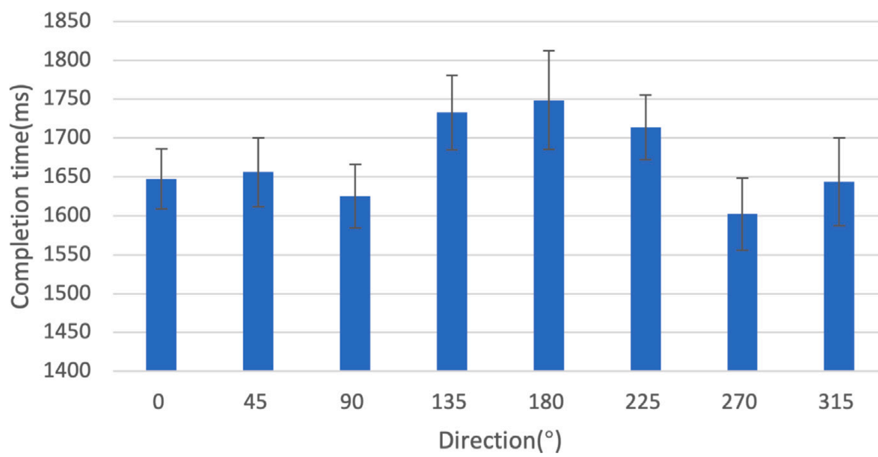

Fig. 16. Completion times for all child participants in different directions. Error bars indicate the standard errors.

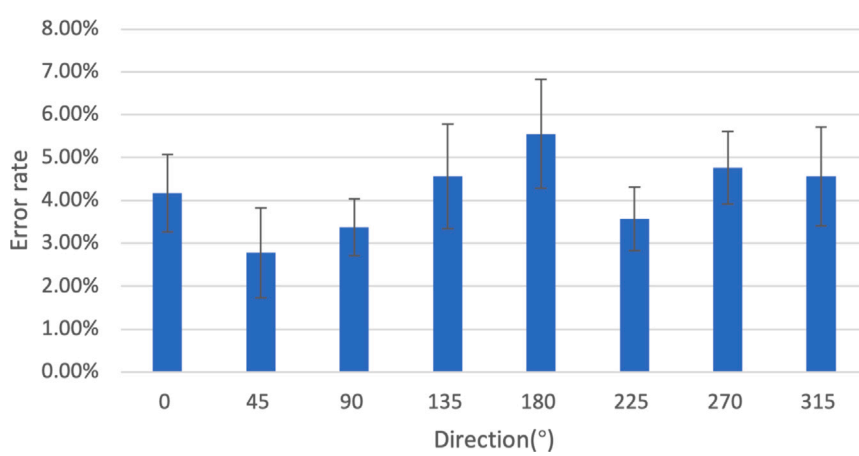

Fig. 17. Error rates for all the child participants in different directions. Error bars indicate the standard errors.

3. The Fitts' law model is well-fitted and can be applied to children of different age groups and different orientations of upper limb movements. The Fitts' law and corresponding metrics, such as ID and throughput, can be used as key references when designing interactive interfaces related to children's upper limb movements. We observed no significant difference in the time to complete the goal-directed task between the two age groups of children in the second experiment. However, completion time and ID showed an approximately linear relationship. Thus, designers can reduce the completion time by carefully designing IDs. Meanwhile, we found a stepwise relationship between the error rate and IDs. When IDs $\leq$ 1.648, the error rate was smaller for all children and lower for participants aged 6–8 years than for those

aged 9–12 years. When IDs ≥ 2.070, error rates increased significantly for all children and participants aged 9–12 years performed better than those aged 6–8 years. Designers should be aware of this difference when designing systems for children of different ages. And our findings provide a reference on ID design from which designers can learn to reasonably design target widths and distances to control IDs to achieve the desired effect.

According to the above design implications, we discuss several possible task scenarios and design prototypes for children's mid-air gesture interactions as follows:

1. Menu components for mid-air gesture interaction

Mid-air gesture interaction has already made a splash in smart homes, virtual reality, and somatic games, and is gradually being embraced by children. As an important part of the somatosensory system, the design of the menu directly affects the ease of operation for child users. Our findings provide guidance for designing more child-friendly menu components for mid-air gesture interaction systems. Here we provide some prototype designs of child-oriented mid-air gesture interaction menus, containing the more common pie, ring, and radial layout menus (Fig. 18(a–c)). According to the aforementioned design implication #1, the maximum number of items that suit for children in such menus is 8. Also, to balance the functionality and efficiency of the menus, we believe that 8 items are appropriate. Fig. 18(d–f) shows the morphing design of these 3 types of regular menus and demonstrates how our findings and design implications provide guidance for the design of such menus. In the future, designers can flexibly apply our findings to design child-oriented mid-air gesture interaction menus based on practical needs. According to the experimental results and the design implication #2, children's accuracy and offset in eight different directions vary widely, so designers can cut the menu into eight unequal parts to accommodate children's arm movement characteristics. For example, children's movement in the left direction has a larger offset and shows instability, while children's arm movement in the right direction is fast but has a high error rate, and the menu design can appropriately increase the area of the left and right button selection area. At the same time, in the pie and ring-shaped menus, the offset of the up and down is small and the accuracy is better, then the selection area in these two directions is allowed to be compressed appropriately according to the actual situation. According to the aforementioned design implication #3, the designer can adjust the distance from the center of the layout circle to the target, in addition to the selection area of the target, to improve the speed and accuracy of children's operation. As illustrated in Fig. 18(f), since children move at a relatively short distance in the left and right directions, the designer can move the targets in these two directions toward the center of the layout circle to facilitate children's hands to capture the targets quickly.

2. Scene design of mid-air gesture interaction games for children

Somatic games are children's favorite gestural interaction systems. The interaction interface is very important, and our study can provide guidance for the design of the game interface. Game designers can better design the size and placement of game props to make it easier for children to choose, or set different levels of difficulty for the game accordingly. For example, in capture or beat games, children need to swing their arms to accurately reach the target object. Experiment 1 yielded data on the maximum range of children's upper limb movements in different directions, error rates, etc., according to which designers can set the directional arrangement of game props. The influence of the number of directions on children's upper limb movement performance is very significant, and the difficulty of the game can be adjusted by increasing or decreasing the direction of the props. The maximum range of the children's arm movements is also an important reference for the layout of props, for example, the movement range of children in the up and down direction is about 43 to 44 cm, and the left and right directions are smaller, about 40 to 42 cm, and these data can be referred to determine the distance of prop placement to improve the ease of operation of the game. In addition, our work provides the relationship between the ID index (determined by width and distance) and children's target capture performance. Designers can draw on the suggestions of the design implications #3 to control the size of the ID index of the game task to achieve the desired game effect, for example, if the distance to the target is 160 pixels and the width of the target is 100 pixels, the ID index is 1.379, which is easier for children. To increase the difficulty, the distance to the target can be increased or the size of the target can be reduced appropriately.

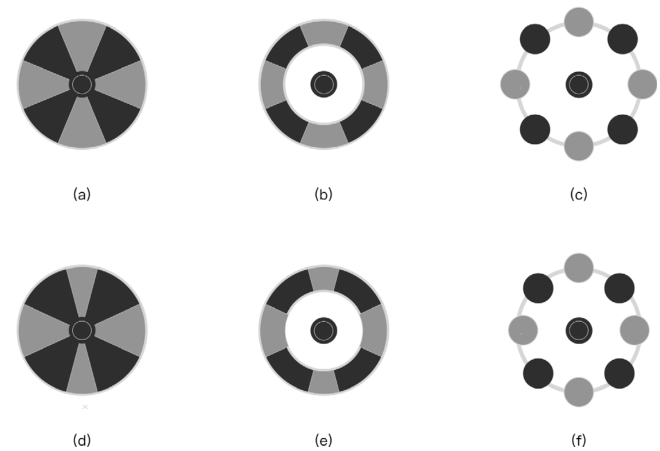

Fig. 18. Pie, ring and radial layout menu prototypes. Note that we highlight each pair of adjacent items in two colors (black and gray) to ensure them distinguishable.

6. Limitations

There are certain limitations in our current study. In our current experiments, the number of children in each age group was small (7 or 10 in each group), which may have biased our results toward individual performances. And, the lack of significant differences between age groups exhibited by some experimental results may also be influenced by small samples. In the future, we will increase the number of children in each age group to further investigate more general differences in performance between different age groups. In addition, the design prototypes provided in this study are preliminary without detailed considerations for specific scenarios. Our current work only discusses how the design implications can be applied to game scenarios. In the future, we plan to design and evaluate more applications and gesture interaction widgets for children and explore diverse application scenarios to demonstrate the applicability of this research.

7. Conclusion and future work

We study children's motor behavior of mid-air upper limb movement in this paper. We found that the 8-item menu offers the most suitable division scheme for children, balancing capacity and accuracy. In spatial orientation, children move most accurately in the up (90°) and down (270°) directions, followed by the due right and downright directions (0° and 315°) with small offsets and low error rates, while the directions near the left side (135°, 180°, and 225°) show poor stability and accuracy. The completion time in different directions can be well predicted by the index of difficulty of target capture tasks in the Fitts' law model. The accuracy of younger children in fine movements, such as 16-item menus or target acquisition tasks with higher indexes of difficulty, is significantly worse than that of older children, although the two age groups perform similarly to other motor tasks. When completing rough movements like waving in 4- or 8-item menus and lower-ID target acquisition tasks, younger children surprisingly perform even faster than older children.

The concerns regarding the spatial constraint design of mid-air gestures of upper limb movement for children include the accuracy in performing upper limb movement and the range of movement. In the design of interactive systems for children, the gesture of waving arms in different directions can follow the results of our first experiment in this paper. We suggest that designers place the most important items in the up-and-down direction, followed by the left–right direction, and lastly in the other directions, in order to enable most children to use upper limb movement accurately and effectively to reach the target. The results of our second experiment inspire the designers to place the selection items within the optimal distance as much as possible with proper sizes.

In the future, we plan to expand our current work in two aspects. First, we will increase the number of children in each age group to generalize the performance differences between the two age groups. Second, we are interested in designing and evaluating more applications of spatial motion-based interface widgets for children to enhance the utility of our research.

CRediT authorship contribution statement

Fei Lyu: Conceptualization, Methodology, Investigation, Formal analysis, Writing – original draft, Writing – review & editing, Project administration. **Huijing Li:** Formal analysis, Data curation, Visualization, Writing – original draft, Writing – review & editing. **Qiang Fu:** Visualization, Writing – original draft, Writing – review & editing. **Yujie Liu:** Software, Data curation, Investigation, Formal analysis, Writing – original draft. **Jin Huang:** Methodology, Writing – original draft, Writing – review & editing. **Zhigang Deng:** Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

The authors would like to thank the anonymous reviewers for their insightful comments. This work was supported by National Natural Science Foundation of China (61972045 and 62172397).

References

Abdul-Rashid, H.M., Kiran, L., Mirrani, M.D., Maraaj, M.N., 2017. CMSWVHG–Control MS windows via hand gesture. In: International Multi-Topic Conference. INMIC, pp. 1–7, URL: https://doi.org/10.1109/INMIC.2017.8289473.

Accot, J., Zhai, S., 1997. Beyond fitts' law: Models for trajectory-based HCI tasks. In: Human Factors in Computing Systems, CHI '97: Looking to the Future, Extended Abstracts. pp. 295–302, URL: https://doi.org/10.1145/258549.258760.

Accot, J., Zhai, S., 2002. More than dotting the i's — Foundations for crossing-based interfaces. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '02, Association for Computing Machinery, pp. 73–80, URL: https://doi.org/10.1145/503376.503390.

Accot, J., Zhai, S., 2003. Refining fitts' law models for bivariate pointing. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '03, Association for Computing Machinery, pp. 193–200, URL: https://doi.org/10.1145/642611.642646.

Adachi, T., Goseki, M., Muratsu, K., Mizoguchi, H., Namatame, M., Sugimoto, M., Kusunoki, F., Yamaguchi, E., Inagaki, S., Takeda, Y., 2013. Human SUGOROKU: Full-body interaction system for students to learn vegetation succession. In: Proceedings of the 12th International Conference on Interaction Design and Children. IDC '13, Association for Computing Machinery, pp. 364–367, URL: https://doi.org/10.1145/2485760.2485830.

Adelson, J.L., Mccoach, D.B., 2010. Measuring the mathematical attitudes of elementary students: The effects of a 4-point or 5-point likert-type scale. Educ. Psychol. Meas. 70 (5), 796–807.

Bachynskyi, M., Müller, J., 2020. Dynamics of aimed mid-air movements. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, New York, NY, USA, pp. 1–12, URL: https://doi.org/10.1145/3313831.3376194.

Bailly, G., Lecolinet, E., Nigay, L., 2008. Flower menus: A new type of marking menu with large menu breadth, within groups and efficient expert mode memorization. In: Proceedings of the Working Conference on Advanced Visual Interfaces. AVI '08, Association for Computing Machinery, pp. 15–22, URL: https://doi.org/10.1145/1385569.1385575.

Balakrishnan, R., MacKenzie, I.S., 1997. Performance differences in the fingers, wrist, and forearm in computer input control. In: Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems. CHI '97, Association for Computing Machinery, New York, NY, USA, pp. 303–310, URL: https://doi.org/10.1145/258549.258764.

Buschek, D., Roppelt, B., Alt, F., 2018. Extending keyboard shortcuts with arm and wrist rotation gestures. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, New York, NY, USA, pp. 1–12, URL: https://doi.org/10.1145/3173574.3173595.

Carvalho, D., Bessa, M., Magalhães, L., 2014. Different interaction paradigms for different user groups: An evaluation regarding content selection. In: Proceedings of the XV International Conference on Human Computer Interaction. In: Interacción '14, Association for Computing Machinery, pp. 1–6, URL: https://doi.org/10.1145/2662253.2662293.

Carvalho, D., Bessa, M., Magalhães, L., Carrapatoso, E., 2015. Performance evaluation of gesture-based interaction between different age groups using fitts' law. In: Proceedings of the XVI International Conference on Human Computer Interaction. In: Interacción '15, Association for Computing Machinery, New York, NY, USA, pp. 1–7, URL: https://doi.org/10.1145/2829875.2829920.

Carvalho, D., Bessa, M., Magalhaes, L., Carrapatoso, E., 2018. Performance evaluation of different age groups for gestural interaction: a case study with Microsoft Kinect and Leap Motion. Univers. Access Inf. Soc. 17 (1), 37–50, URL: https://doi.org/10.1007/s10209-016-0518-4.

Cheema, N., Frey-Law, L.A., Naderi, K., Lehtinen, J., Slusallek, P., Hämäläinen, P., 2020. Predicting mid-air interaction movements and fatigue using deep reinforcement learning. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, pp. 1–13, URL: https://doi.org/10.1145/3313831.3376701.

Chen, Z., Chen, Y.-P., Shaw, A., Aloba, A., Antonenko, P., Ruiz, J., Anthony, L., 2020. Examining the link between children's cognitive development and touchscreen interaction patterns. In: Proceedings of the 2020 International Conference on Multimodal Interaction. Association for Computing Machinery, New York, NY, USA, pp. 635–639, URL: https://doi.org/10.1145/3382507.3418841.

Connell, S., Kuo, P.-Y., Liu, L., Piper, A.M., 2013. A wizard-of-oz elicitation study examining child-defined gestures with a whole-body interface. In: Proceedings of the 12th International Conference on Interaction Design and Children. IDC '13, Association for Computing Machinery, pp. 277–280, URL: https://doi.org/10.1145/2485760.2485823.

Feldman, R., 2008. Development Across the Life Span: Pearson New International Edition. Pearson Schweiz Ag.

Fitts, P.M., 1954. The information capacity of the human motor system in controlling the amplitude of movement. J. Exp. Psychol. 47 (6), 381–391, URL: https://doi.org/10.1037/h0055392.

Gang, R., O'Neill, E., 2012. 3D marking menu selection with freehand gestures. IEEE 61–68.

Garcia-Sanjuan, F., Nacher, V., Jaen, J., 2016. MarkAirs: Are children ready for marker-based mid-air manipulations? In: Proceedings of the 9th Nordic Conference on Human-Computer Interaction. NordiCHI '16, Association for Computing Machinery, New York, NY, USA, pp. 1–8, URL: https://doi.org/10.1145/2971485.2971517.

Grossman, T., Balakrishnan, R., 2004. Pointing at trivariate targets in 3D environments. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, New York, NY, USA, pp. 447–454, URL: https://doi.org/10.1145/985692.985749.

Hourcade, J.P., 2022. Child–Computer Interaction.

Jang, S., Stuerzlinger, W., Ambike, S., Ramani, K., 2017. Modeling cumulative arm fatigue in mid-air interaction based on perceived exertion and kinetics of arm motion. In: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, New York, NY, USA, pp. 3328–3339, URL: https://doi.org/10.1145/3025453.3025523.

Lu, F., Tian, F., Jiang, Y., Cao, X., Luo, W., Li, G., Zhang, X., Dai, G., Wang, H., 2011. ShadowStory: Creative and collaborative digital storytelling inspired by cultural heritage. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, pp. 1919–1928, URL: https://doi.org/10.1145/1978942.1979221.

Lyu, F., Tian, F., Feng, W., Cao, X., Zhang, X.L., Dai, G., Wang, H., 2017. EnseWing: Creating an instrumental ensemble playing experience for children with limited music training. In: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, New York, NY, USA, pp. 4326–4330, URL: https://doi.org/10.1145/3025453.3025583.

Lyu, F., Xi, R., Han, Y., Liu, Y., 2018. MagicMark: a marking menu using 2D direction and 3D depth information. Sci. China Inf. Sci. 61 (6), 1–3.

MacKenzie, I.S., 2015. User studies and usability evaluations: From research to products. In: Proceedings of the 41st Graphics Interface Conference. GI '15, Canadian Information Processing Society, CAN, pp. 1–8.

MacKenzie, I.S., Buxton, W., 1992. Extending fitts' law to two-dimensional tasks. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '92, Association for Computing Machinery, New York, NY, USA, pp. 219–226, URL: https://doi.org/10.1145/142750.142794.

MacKenzie, I.S., Isokoski, P., 2008. Fitts' throughput and the speed-accuracy tradeoff. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '08, Association for Computing Machinery, New York, NY, USA, pp. 1633–1636, URL: https://doi.org/10.1145/1357054.1357308.

MacKenzie, I.S., Kauppinen, T., Silfverberg, M., 2001. Accuracy measures for evaluating computer pointing devices. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '01, Association for Computing Machinery, pp. 9–16, URL: https://doi.org/10.1145/365024.365028.

Mine, M.R., Brooks, F.P., Sequin, C.H., 1997. Moving objects in space: Exploiting proprioception in virtual-environment interaction. In: Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques. SIGGRAPH '97, ACM Press/Addison-Wesley Publishing Co., USA, pp. 19–26, URL: https://doi.org/10.1145/258734.258747.

Moser, C., Tscheligi, M., 2015. Physics-based gaming: Exploring touch vs. Mid-air gesture input. In: Proceedings of the 14th International Conference on Interaction Design and Children. IDC '15, Association for Computing Machinery, New York, NY, USA, pp. 291–294. http://dx.doi.org/10.1145/2771839.2771899, URL: https://doi.org/10.1145/2771839.2771899.

Muangmoon, O.-o., Rattanakhum, M., Sureephong, P., 2016. Game menu navigation for the elderly using non-tactile gesture interaction: a pilot study. In: Proceedings of the International Convention on Rehabilitation Engineering & Assistive Technology. pp. 1–4.

Nacenta, M.A., Kamber, Y., Qiang, Y., Kristensson, P.O., 2013. Memorability of pre-designed and user-defined gesture sets. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '13, Association for Computing Machinery, New York, NY, USA, pp. 1099–1108, URL: https://doi.org/10.1145/2470654.2466142.

Nam, H.S., Lee, W.H., Seo, H.G., Kim, Y.J., Bang, M.S., Kim, S., 2019. Inertial measurement unit based upper extremity motion characterization for action research arm test and activities of daily living. Sensors 19 (8), 1782.

Ni, T., Bowman, D.A., North, C., Mcmahan, R.P., 2011. Design and evaluation of freehand menu selection interfaces using tilt and pinch gestures. Int. J. Hum.-Comput. Stud. 69 (9), 551–562.

Ni, T., McMahan, R.P., Bowman, D.A., 2008. Tech-note: RapMenu: Remote menu selection using freehand gestural input. In: Proceedings of the 2008 IEEE Symposium on 3D User Interfaces. In: 3DUI '08, IEEE Computer Society, USA, pp. 55–58, URL: https://doi.org/10.1109/3DUI.2008.4476592.

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

Pang, X.Y., Guo, R.Z., Yao, N.L., Jia-Lin, Y.U., Shi-Xian, Y.U., Wang, C., Gao, Z.F., Psychology, D.O., University, Z., 2014. Human factor studies on gestural interaction:Past,present,and future. Chin. J. Appl. Psychol. 243–251.

Pereira, A., Wachs, J.P., Park, K., Rempel, D., 2015. A user-developed 3-D hand gesture set for human–computer interaction. Hum. Factors 57 (4), 607–621, URL: https://doi.org/10.1177/0018720814559307.

Royeen, C.B., 1985. Adaptation of likert scaling for use with children. Otjr Occup. Particip. Health (1), 59–69, URL: https://doi.org/10.1177/153944928500500104.

Rubegni, E., Gentile, V., Malizia, A., Sorce, S., Kargas, N., 2019. Child-display interaction: Exploring avatar-based touchless gestural interfaces. In: Proceedings of the 8th ACM International Symposium on Pervasive Displays. PerDis '19, Association for Computing Machinery, New York, NY, USA, pp. 1–7, URL: https://doi.org/10.1145/3321335.3324942.

Ruiz-Rodriguez, A., Martinez-Garcia, A.I., Caro, K., 2019. Gesture-based video games to support fine-motor coordination skills of children with autism. In: Proceedings of the 18th ACM International Conference on Interaction Design and Children. IDC '19, Association for Computing Machinery, New York, NY, USA, pp. 610–615, URL: https://doi.org/10.1145/3311927.3325310.

Sanchez, V.L.A., Cruz, O.I.I., Solorza, E.A.A., Monroy, I.A.E., Caro, K., Castro, L.A., 2017. BeeSmart: A gesture-based videogame to support literacy and eye-hand coordination of children with down syndrome. In: International Conference on Games and Learning Alliance. Springer, pp. 43–53.

Tian, F., Lyu, F., Zhang, X., Ren, X., Wang, H., 2017. An empirical study on the interaction capability of arm stretching. Int. J. Hum.–Comput. Interact. 33 (7), 565–575.

Vatavu, R.-D., Anthony, L., Wobbrock, J.O., 2013. Relative accuracy measures for stroke gestures. In: Proceedings of the 15th ACM on International Conference on Multimodal Interaction. ICMI '13, Association for Computing Machinery, pp. 279–286, URL: https://doi.org/10.1145/2522848.2522875.

Vatavu, R.-D., Cramariuc, G., Schipor, D.M., 2015. Touch interaction for children aged 3 to 6 years: Experimental findings and relationship to motor skills. Int. J. Hum.-Comput. Stud. 74, 54–76.

Wittorf, M.L., Jakobsen, M.R., 2016. Eliciting mid-air gestures for wall-display interaction. In: Proceedings of the 9th Nordic Conference on Human-Computer Interaction. NordiCHI '16, Association for Computing Machinery, New York, NY, USA, pp. 1–4, URL: https://doi.org/10.1145/2971485.2971503.

Yan, J.H., Thomas, J.R., Stelmach, G.E., Thomas, K.T., 2000. Developmental features of rapid aiming arm movements across the lifespan. J. Motor Behav. 32 (2), 121–140.

Zhai, S., Kong, J., Ren, X., 2004. Speed–accuracy tradeoff in Fitts' law tasks—on the equivalency of actual and nominal pointing precision. Int. J. Hum.-Comput. Stud. 61 (6), 823–856.