Effects of Continuous Auditory Feedback on Drawing Trajectory-Based Finger Gestures

BoYu Gao, *Student Member, IEEE*, HyungSeok Kim, *Member, IEEE*, Hasup Lee, Jooyoung Lee, and Jee-In Kim, *Member, IEEE*

***Abstract*—The well-known "fat finger" issue limits the interaction performance of trajectory-based finger gestures. To alleviate this issue, this work focuses on the possibility of using additional continuous auditory feedback to assist trajectory-based finger gestures. First, the experiment validated that, with the visual feedback only, the bare fingertip led to more errors in drawing of intersectional points, endpoints of closed gestures, and gestural length and shape variability compared to when the finger-attached pen was used. Then, we designed different types of auditory feedback (discrete beep, static, gradual) to provide additional information on the spatial relationship between finger-contact point and the endpoints or intersections of predefined gestures. An experiment that evaluates the effects of individual or combination of designed auditory feedback on trajectory-based finger gestures was conducted. These results show a few differences between them. However, a combination of gradual (amplitude and frequency) continuous sound and beep reached the highest drawing accuracy for trajectory-based finger gestures, which is similar to that of a finger-attached pen. This research offers insights and implications for the future design of continuous auditory feedback on small touchscreens.**

***Index Terms*—Continuous auditory feedback, small touchscreens, trajectory-based finger gestures, visual occlusion.**

I. Introduction

With the increasing adoption of touchscreen devices, fingers and pens have become important interaction modalities in gestural touch interfaces. Generally, people are used to utilizing the finger modality on touch screens thanks to its simplicity and a convenience for tasks, e.g., target acquisition [6]–[9], typing [16], item-dragging, and scrolling [5]. However, a critical challenge in applying finger modality on touch screens is its inaccuracy, especially relative to smaller-sized targets [4], due to the well-known "fat finger" problem [2] (visual occlusion) and *perceived input point* model [9]. The visual feedback is therefore not feasible all the time when utilizing the finger modality on touch screens. However, most of these interactions are cognitively mastered in response to information that is visually perceived [1], i.e., "what you see is what you get."

Manuscript received June 9, 2017; revised December 14, 2017 and May 16, 2018; accepted June 9, 2018. Date of publication July 10, 2018; date of current version November 13, 2018. This work is supported by the Bio-Synergy Research Project under Grant 2013M3A9C4078140 of the ministry of science, ICT and future planning through the National Research Foundation. This paper was recommended by Associate Editor H. Liu. *(Corresponding author: HyungSeok Kim.)*

The authors are with the Department of Software, Konkuk University, Seoul 05029, South Korea (e-mail: yuye@konkuk.ac.kr; hyuskim@konkuk.ac.kr; hasups@gmail.com; ljy1201@konkuk.ac.kr; jnkm@konkuk.ac.kr).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/THMS.2018.2850329

Prior representative approaches [6]–[9] [11] have been proposed for alleviating the issue of "fat finger" by compensating additional visual feedback. For example, Albinsson *et al.* [6] proposed *CrossKeys* and *PrecisionHandle* to allow pixel level pointing in a fast and efficient manner. More recently, *FFitts law* [16] has been proposed to more accurately model finger touch for target acquisition and typing on small touchscreens, compared to *Fitts' law* [17]. However, note that these approaches [6]–[9], [11], [16] mainly focused on modeling discrete finger-touch interactions, e.g., typing [16] and target acquisition [6], [11]. Subsequent works, essentially concerned on pen-based steering tasks (also known as trajectory-based tasks), have been done to build models for steering through corners [21], steering within above-the-surface interaction layers [22]. These approaches relied on visual feedback, as Cockburn *et al.* [23, p. 1] stated *the visual feedback only simulates human visual system, leaving the powerful auditory and tactile senses redundant.*

The use of multimodal feedback for pointing, selecting, and steering-tasks with mouse or pen modalities have been investigated [12], [13], [23], [33]. For example, the multimodal feedback (nonspeech audio, tactile via vibration, and pseudohaptic "sticky" feedback) modes helped reduce the targeting time for small and discretely located targets [23]. Brewster [13] found that adding-sounds-to-buttons made it significantly easier for users to select, and overcome the lack of screen space on mobiles. However, these works mainly focused on the discrete actions by mapping multimodal alarm feedback.

Later, Sun *et al.* [33] investigated the effects of multimodal error feedback on human performance in steering-tasks with stylus, where the accuracy performance other than the movement time was improved with additional stimulus (auditory, haptic). The error feedback was defined as the tracking and trajectory errors are happening, which is a kind of an event notification.

Andersen *et al.* [12] conducted a series of experiments to explore the effectiveness of mapping the perceptual auditory feedback with freeform pen-gestures. Results showed that performance gains can be achieved with the use of stated auditory feedback after gesture completion. In particular, it indicated that the performances regarding the drawing of closed strokes and the intersection points between pairs of lines were influenced by both visual and auditory feedback.

It is important to note that these used input modalities with auditory alarm feedback in target acquisition and steering tasks were mouse cursor and pen stylus, which have not suffered from the issues of "fat finger." Later, Cockburn *et al.* [4] investigated the pointing, tapping, and dragging tasks with finger, mouse, and stylus modalities, and found that finger resulted in slower time than other two modalities in dragging tasks, and *dragging errors were low in all conditions*.

Meanwhile, Tu *et al.* [3] studied the similarities and differences between finger and pen modalities on drawing freeform gestures on touch screens, and found that drawing geometric shape features (*aperture of closed strokes, corner shape distance, and intersecting points deviation*) with finger was more difficult than with pen, these features are important to the stroke recognition [26].

Despite the "fat finger" issue, the trajectory-based finger gesture becomes pervasive in various applications, for example, navigating in circular menus [29], calligraphy practice [33], and driving simulation [34]. However, little work has been done to understand effects of multimodal feedback [44] on trajectory-based finger gestures. Zhao *et al.* [29, p. 1] stated that "the technology required to generate tactile feedback is typically complex. Auditory feedback would seem to be a more promising option, especially given existing technological support in mobile devices such as phones and media players." Therefore, the well-known "eyes-free" menu selection technique using reactive audio feedback was proposed with trajectory-based finger gestures. However, these mentioned types of auditory feedback [13], [23], [29], [33], [41] mainly provide the alarm or sonification of movement error.

A trajectory-based finger gesture, however, is a continuous action for which continuous feedback is required for the achievement of natural and intuitive interactions [14], [39]. For example, the continuous auditory channel can provide additional persistent and informative feedback for finger touch gestures [15], and improving accuracy of planned movements, regularity, and reliability of guidance [35], [36]. Sigrist *et al.* [42] presented a review of augmented multimodal feedback for motor learning, in particular, for complex motor task learning [43], and revealed the potential of concurrent auditory feedback. More recently, Boyer *et al.* [35] investigated three types of continuous auditory feedback in visuo-manual tracking task, and found them beneficial for movement training. However, this visuo-manual task traces the trajectories of stylus movement, which have not suffered from "fat-finger" problem.

Prior works [15], [19] have shown advantages of mapping continuous auditory feedback with finger gestures for visually impaired users, compared with haptic guidance. However, less is known on the effects of continuous and discrete types of auditory feedback on trajectory-based finger gestures.

Another pilot study, where beep sound feedback was used to provide occluded spatial information with the trajectory-based finger gestures [10], was conducted. Discrete beep sound, such as playing a sound when the user's finger crosses on the boundary area, helped the users to become aware of the drawing state of the intersectional points and the endpoints of predefined gesture patterns. However, the beep sound has been primarily limited to indicate changes of state in the interface. It was insufficient for trajectory-based gestures as the gestural position is continuously changing. For example, in the real situation, the notification for a park distance control will be given when a car is getting closer to the obstacle over time. Generally, the car provides a duration of continuous warning to drivers, instead of a discrete alarm notification.

To our knowledge, less is known on addressing the role of continuous auditory feedback in drawing trajectory-based finger gestures. We argue that the trajectory-based finger gestures require more continuous and onsite auditory feedback, as the user is continuously dragging with the use of finger modality on such small touchscreens. Toward the end, the following specific

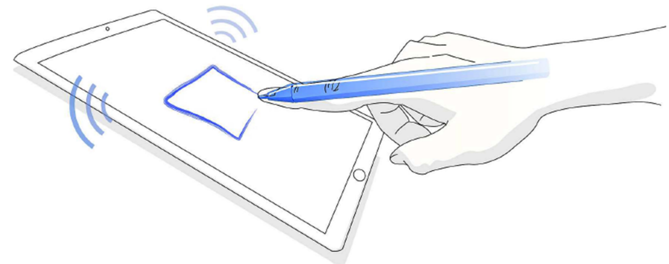


Fig. 1. Conceptual image illustrates the focus of this work that utilizes continuous auditory feedback to compensate for drawing trajectory-based finger gestures, which is similar to a finger-attached pen.

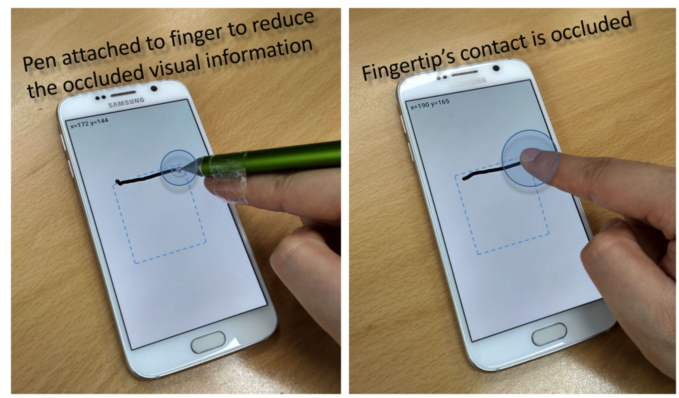


Fig. 2. Evaluating the effects of concurrent visual feedback on trajectory-based gesture drawing. The pen attached to index finger is to reduce the occluded visual feedback, while fingertip's contact area is occluded.

questions are proposed to study the effects of continuous auditory feedback on trajectory-based finger gestures, so that users may provoke the experience of drawing with pen or stylus (see Fig. 1).

1) What types of spatial visual information are occluded when drawing the trajectory-based finger gestures?
2) What types of perceptual auditory feedback can be used to map with the occluded visual information?
3) How does continuously perceptual auditory feedback enhance the performance of trajectory-based finger gestures?

In this work, the index fingertip and finger-attached pen conditions were carefully controlled to validate the effects of concurrent visual feedback on drawing trajectory-based gestures (see Fig. 2). The drawing of local geometric features of these gestures using finger modality, such as intersectional points of line segments and endpoints of closed strokes [3], [10], [12], were error-prone, compared with the finger-attached pen. In addition, the global geometric features such as gestural length difference and shape variability [27] were larger with index fingertip than with a finger-attached pen condition, most likely due to the occluded visual feedback (see Fig. 4). The results in the trajectory-based tasks reinforced the findings of prior works [3], [12] that were conducted in the freeform drawing tasks.

From the result of the first experiment, three types of auditory feedback (discrete, static, and gradual) were therefore designed to provide different levels of temporal information for local and global geometric features. The areas of endpoints and intersections of template are coupled with different types of auditory feedback, as shown in Fig. 6. The spatial relation between fingertip-contact point and the endpoint or intersection point of template gesture was established using continuous auditory parameters (amplitude and frequency) (see Fig. 7). The feedback conditions such as an individual or combination of the auditory feedback were controlled as the independent variable. A comparative experiment that evaluates the performance of trajectory-based finger gestures under these feedback conditions was conducted. Results showed few differences between them. The gradual continuous auditory feedback gave better drawing accuracy than discrete beep sound on all geometric measures (see Fig. 8 and Table II). The consistence analysis of accuracy among these feedback conditions across the total number of trials was observed and showed that trajectory-based finger drawn gesture with combination of the beep and the gradual auditory feedback was similar to finger-attached pen drawn (see Fig. 9).

The rest of paper is divided into six sections. Section II describes the evaluation of the effects of visual feedback on trajectory-based drawing tasks. The auditory feedback design is illustrated in Section III. Section IV shows the effects of different types of auditory feedbacks on finger modality in the same task. Sections V and VI conclude this work with the discussion and its potential future work.

II. Experiment 1: Effects of Concurrent Visual Feedback on Trajectory Drawing

From the previous observations, the impacts of visually occluded feedback on accuracy of finger interaction in the trajectory-based gesture task were first evaluated. The trajectory -based drawing task was chosen as the test bed, as it involves a visually heavy workload. The visually occluded information was carefully designed as an independent variable, by utilizing very small-diameter pen that could be attached to the index finger of dominant hand, as shown in Fig. 2. The reason why we employed a finger-attached pen is to alleviate the confounding effect on the pen-grip method [18]. The formulated hypothesis is that a finger-attached pen could give more visual feedback, thereby allowing the participants to achieve better performance of drawing trajectory-based gestures.

A. Participants

Fifteen volunteers comprising of six females and nine males, who were aged from 20 to 35 (average: 26.5, SD 3.2), with backgrounds in the fields of art design, computer science, and international finance, participated in the experiment. Twelve of the participants are right-hand dominant, while the other three are left-hand dominant. Four of the participants had previously drawn on touchscreens with either the pen stylus or the finger; two had no experience with both drawing implements, while nine of the participants had prior experience regarding the use of both modalities.

B. Apparatus

An android application was developed for the experimental trajectory-based drawing task in this study, as shown in Fig. 2. A common Galaxy S6 mobile phone, which can support both trajectory-based finger- and pen-drawn gestures, with the following specifications: a relatively small touchscreen 5.1 in and 1440 pixels × 2560 pixels (577 ppi pixel density). While the capabilities of individual smartphones are typically considered as different, for this experiment, it was assumed that the latency difference between the interface response and the user input is in the accepted error-tolerance level under each experimental condition. The pen that was used in this experiment is the *Adonit Jot* [25], a precision degree of which is supposed to be similar to a 0.5 mm ball-point pen. As different pen-nib diameters can also influence the pen-drawing performance, for example, Annett *et al.* [24] designed four different nib-diameter sizes as a control factor, and the results showed a lower accuracy for the larger diameters, as the larger diameters limited the provision of visual feedback. In addition, these touch events were registered and recorded by developed program for the subsequent data analysis. These touch events include the x and y coordinates, and a timestamp of each event.

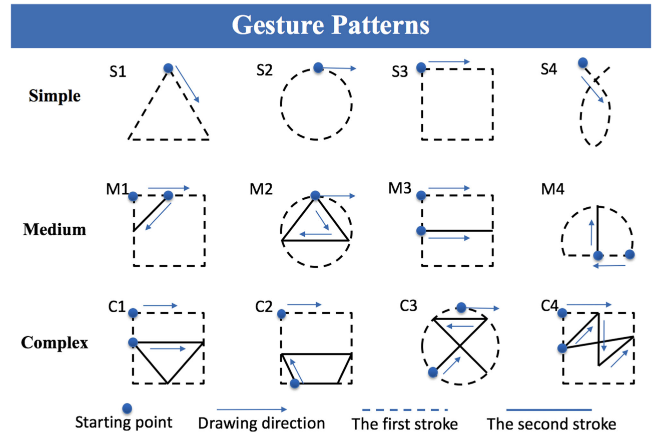


Fig. 3. Gesture patterns used in this experiment. The details of tracing the gesture patterns, e.g., there are two strokes in M1, the dotted line represents the first stroke, and the solid dot represents the starting point. The arrow shows the drawing direction. After the first stroke, participants then continue to the second stroke that is marked with the solid line.

TABLE I
Measures Selection

Measures	Categories	Descriptions
1.Gesture-Completion Time	Basic feature	1. The duration between the start and the end of the drawn stroke.
2. Aperture Distance	Local shape features	2. The Euclidean distance between the starting and ending positions of the drawn stroke.
3. Intersection-Point Deviation	Local shape features	3. The Euclidean distance between the drawn gesture and the template for the intersection point.
4. Length Difference	Global shape features	4. Length difference between the drawn gesture and the template.
5. Shape Variability	Global shape features	5. The standard deviation of the distances between the points of the drawn and the template.

C. Gesture Patterns Design

According to the modeling user performances of steering and freeform drawing tasks [3], [20], the design of gestures that mainly consist of corners, multiple line segments, and circular strokes was considered. Some of the gesture patterns were adopted from existing works [3], [10] (S1, S2, S3, and S4 were from [3], M1, C1, and C4 were from [10]), but the newly

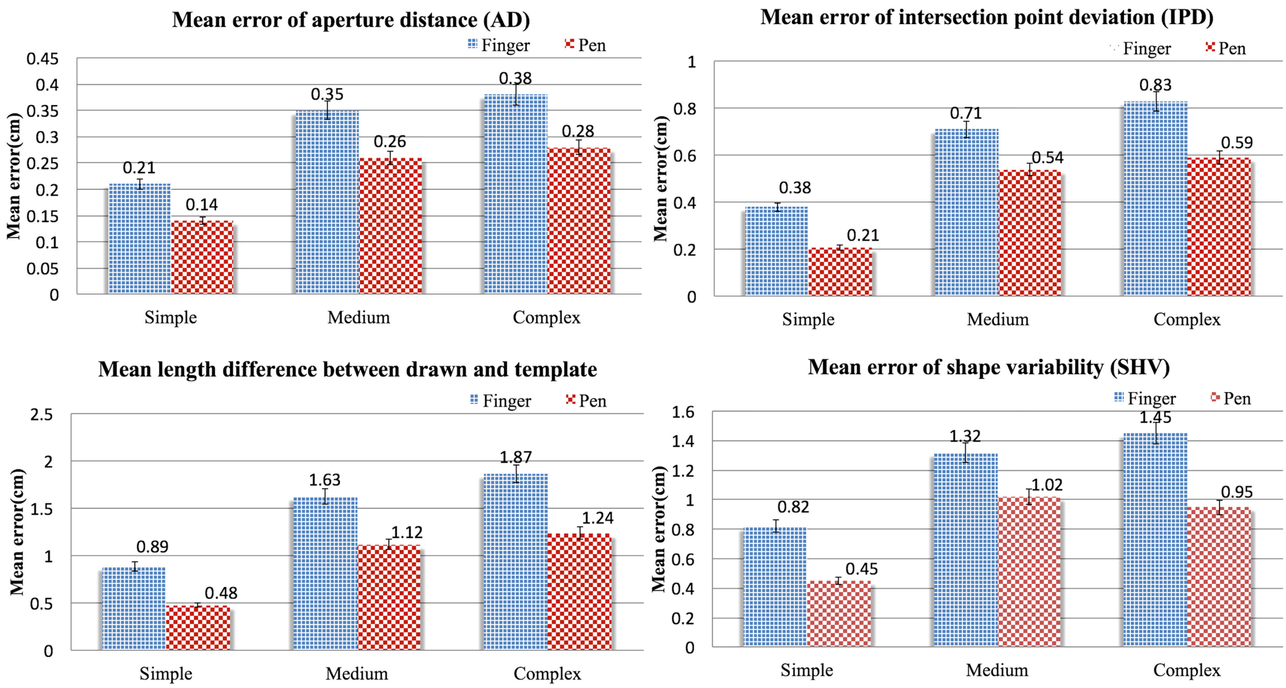


Fig. 4. The effects of visual feedback on geometric measures between finger and finger-attached pen. Results show that visually occluded information contributes to low accuracy of finger modality in the trajectory-based task, while finger-attached pen gives better accuracy on these four measures. The significant differences on each measure between finger and finger-attached pen were found with repeated measures ANOVA, if we ignored the gesture complexity. Error bar represents 95% confidence interval. (Left bar represents the performance of finger.)

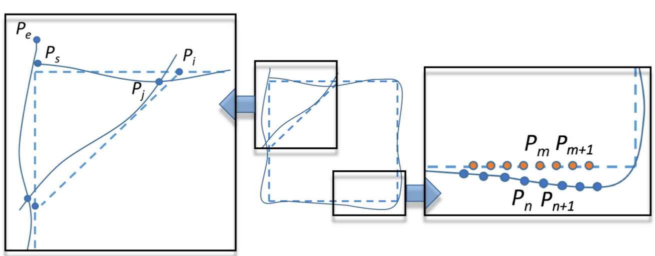


Fig. 5. Calculating the local and global measures, P_s and P_e represent the starting and ending point of the gesture, respectively, while the difference represents the AD. P_i and P_j represent the intersectional point from the template and drawn gesture, the difference represents the IPD. P_m and P_n represent the sample points from the template and drawn gesture, respectively, and the difference between each corresponding sample represents the SHV.

designed gestures that were numbered M2, M3, M4, C2, and C3 have been added for this study. They were classified into three levels of complexity according to the number of corners, intersections, and line segments [20], as shown in Fig. 3. Although the gesture set could not represent all types of gestural features, the typical features such as intersections between line segments, endpoints of closed strokes were highlighted in this gesture set.

D. *Visual Geometric Features Selection*

Prior works [3], [12], [27] have shown the difficulty of accurately drawing a series of geometric features for the closed strokes and the intersections between pairs of line segments with finger in free-from stroke drawing tasks [3], and the visual and auditory feedback impacted the drawing performances using a pen modality [12]. The *Euclidean Distance* between the starting and ending positions of the closed stroke indicates an ability to return to the same point under feedback conditions [10], [12]. This distance reflects the same type of ability that crosses "o" shapes and "b" shapes in handwriting research [26]. The extent to which the relative positions of the intersection points in the drawn gesture were changed from the template is another indication of the shape difference for the evaluation of the drawing performance [3].

Globally, the finger-drawn gestures are often bigger than pen-drawn [3]. The shape variability [27] was used to measure the average global difference between a user-drawn gesture and a template. The time performance is the basic kinematic feature [3]. These mentioned visual features are the indications of ability to draw the gestures with finger and pen modalities under different feedback conditions. The descriptions of each feature are shown in Table I.

E. *Tasks and Procedures*

We obtained approval from the ethics committee at Konkuk University to run this study. Each participant was required to sign the consent form before the experiment. Then, the participants were provided with an informed explanation of each procedure. In the training phase, the practices with the trajectory-based drawing interface using the index finger of dominant hand and a finger-attached pen for 20 trials were given to each participant, respectively. For the experimental stage, the participants were asked to follow the template to trace each of 12 gesture patterns as accurately as possible in a random with counterbalanced order. In total, they were asked to draw all gesture patterns with index finger and a finger-attached pen for four rounds. After each round, each participant was given a 10-min break.

The total number of drawing trials of this experiment was 1320 (15 subjects $\times$ 2 conditions $\times$ 12 gestures $\times$ 4 rounds $= 1320$).

F. Results and Analysis

Within-subjects repeated measures ANOVA was used to evaluate the effects of visually occluded information on drawing gesture tasks. The complexity of gesture pattern served as the other independent variable. Tukey's honestly significant difference (HSD) posthoc test was employed for multiple comparisons when significant difference was found.

Average gesture-completion time: The average gesture-completion time is equal to the sum of the completion time over all of the 12 gestures. The average gesture-completion time with finger and finger-attached pen was 3.64 s and 4.12 s, respectively, across all trials. Significant differences for both the drawing implements ($F_{1,14} = 4.38$, $p < 0.05$) and the gestural complexity ($F_{2,28} = 78.6, p < 0.001$) on the gesture-completion time were found. An interaction effect occurred between the gestural complexity and the drawing implements with respect to the gesture-completion time ($F_{2,28} = 7.64, p < 0.01$). Finger performed faster than finger-attached pen in drawing medium (4.25 s versus 4.74 s) and complex (4.72 s versus 5.24 s) levels of gestures.

Aperture distance (AD): We calculated the mean error of all of the closed gestures for the AD as follows:

$$AD = \frac{1}{k}\sum_{k=1}^{k}(\overline{\overline{P_e - P_s}}) \tag{1}$$

where *k* represents the number of closed gestures, in total 11 closed gestures, except for S4; P_s and P_e represent the starting and ending points of the closed stroke, as shown in Fig. 5. The average AD of finger and finger-attached pen at each level of complexity is shown in Fig. 4 (left top). As expected, the finger-attached pen conceivably performed better than the sole use of the finger on the AD of the closed strokes (finger-attached pen: 0.23 cm versus finger: 0.31 cm, $F_{1,14} = 4.26$, $p < 0.05$) since the visual feedback was occluded for finger modality. The participants could accurately draw the closed strokes using the finger-attached pen as it was easier to see the exact location of the drawing tips. However, a significant difference was not evident on the AD for the gesture complexity ($F_{2,28} = 0.26$, $p = 0.063$). There was no interaction effect between the gestural complexity and the drawing implements with respect to the AD ($F_{2,28} = 0.41$, $p = 0.07$).

Intersection-point deviation (IPD): The mean error of the IPD was summed for the nine gestures of S4, M1, M2, M3, M4, C1, C2, C3, and C4 all of which have intersectional points. The mean IPD of one gesture is represented as follows:

$$\mathrm{IPD} = \frac{1}{n}\sqrt{\sum_{i,j=1}^{n}(P_i - P_j)^2} \tag{2}$$

where *n* represents the number of intersectional points in one of the nine gestures, P_{i} and P_j represent the intersection point from template and drawn gesture, respectively. This Euclidean distance difference between P_i and P_j is the deviation of the intersection point, as shown in Fig. 5.

The average IPD of finger and finger-attached pen in each level of complexity is shown in Fig. 4 (top right), where finger-attached pen performed better on the IPD ($F_{1,14} = 9.61$, $p < 0.01$). Finger-attached pen input resulted in the IPD with 0.44 cm and finger input produced the IPD with 0.64 cm, probably due to the occluded visual information. Different from the AD, a significant difference was found for the gestural complexity on the IPD ($F_{2,28} = 13.46, p < 0.001$). Posthoc test revealed significant differences between all of complexity level pairs at $p < 0.05$ level, with the exception of the result between medium and complex levels. An interaction effect between the gestural complexity and the drawing implements for the IPD was found ($F_{2,28} = 5.63$, $p < 0.01$). Finger led to more errors in drawing three levels of gesture patterns (see Fig. 4). There were smaller errors when drawing simple gestures than drawing medium and complex gestures, but no difference between them.

Length difference (LD): To calculate the LD between the drawn stroke and the template, the gestural length that consists of a sequence of sample points was formulated. The Euclidean distance L_k of one gesture is shown in (3), where M_p is the number of sample points, P_{n+1} and P_n represent two of successive sample points. The LD of a single gesture is represented in (4), where *L* is the length of the template

$$L_k = \sum_{n=1}^{M_p - 1}\sqrt{[P(n+1)_x - P(n)_x]^2 + [P(n+1)_y - P(n)_y]^2} \tag{3}$$

$$L_{\text{diff}} = L_k - L_{\text{temp}}. \tag{4}$$

Fig. 4 (left bottom) shows the mean of LD using finger and finger-attached pen for three complexity levels of gestures. As expected, the LD of finger drawn gesture was bigger than pen drawn (1.46 cm versus 1.15 cm, $F_{1,14} = 5.12$, $p < 0.01$), even in the case of tracing the templates.

There was a significant difference for the gestural complexity on the LD ($F_{2,28} = 13.46$, $p < 0.001$). Posthoc test showed the same trend as the IPD for LD ($p < 0.05$), when the level of gestural complexity increased, more LD errors occurred, but no difference between medium and complex gestures was found. An interaction effect between the gestural complexity and the drawing implements for the LD was found ($F_{2,28} = 6.31$, $p < 0.01$). Finger drawn gestures were bigger than finger-attached pen drawn gestures in three levels of gestural complexity (see Fig. 4).

Shape variability (SHV): the SHV is a computation of the standard deviation of the distance between the drawn-gesture point and the template-gesture point [27]. The sampling template is divided into several equivalent line segments, then the differences between drawn-line segments and the template are compared to obtain the SHV value.

$$\mathrm{SHV} = \sqrt{\frac{1}{k}\sum_{n,m=1}^{k}(P_n - P_m)^2}. \tag{5}$$

Equation (5) describes the calculated variability of a single gesture, where *n* represents the number of sample points from each gesture, P_n and P_m represent the drawn-gesture point and the template-gesture point respectively. Large SHV values indicate that errors are larger for some parts of the gesture (see Fig. 5).

The SHV value for finger modality resulted in a larger error across the gestural paths than for finger-attached pen (1.19 cm versus 0.81 cm, $F_{1,14} = 6.08$, $p < 0.01$). A significant difference was not evident for the gestural complexity on the SHV

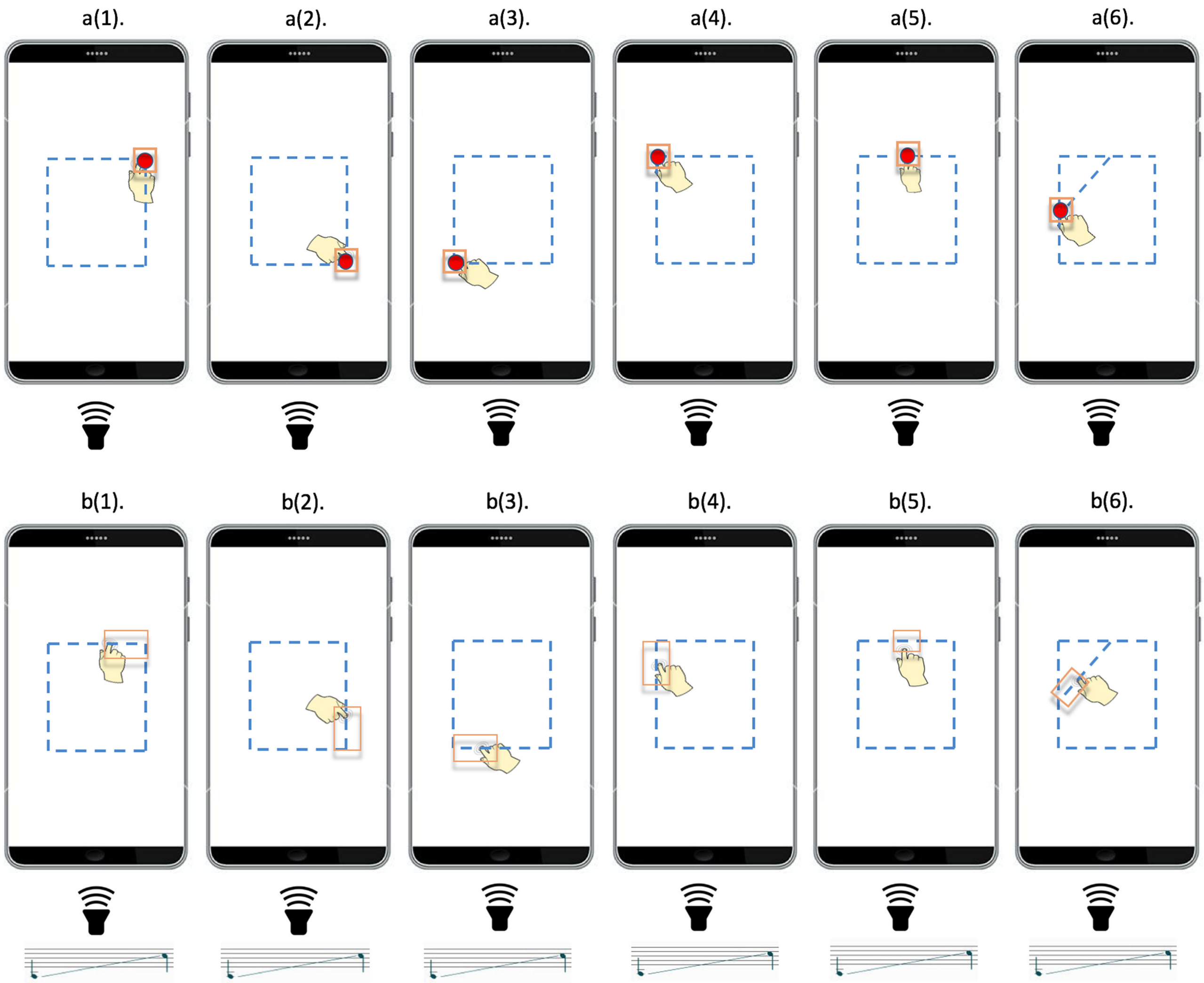


Fig. 6. The above figures illustrate one example of drawing gestures with beep and continuous auditory feedback. In a(1)–a(6), the small rectangle contained endpoint represents the beep sound area. When the fingertip is tracing across the area of endpoint or intersection of template, the beep sound is given. In b(1)–b(6), the rectangle area represents the continuous auditory feedback area where we map the frequency, amplitude with the normalized distance (see Fig. 7). When the fingertip is tracing in the feedback area, the continuous auditory feedback that provides frequency and/or amplitude assists drawing trajectory-based finger gestures. The details of designing auditory feedback are described in Section III.

($F_{2,28} = 0.46$, $p = 0.074$). However, there was an interaction effect between the gestural complexity and the drawing implements on the SHV ($F_{2,28} = 5.16$, $p < 0.05$). More errors occurred when drawing medium and complex gesture patterns using both drawing modalities than drawing simple gestures, but no difference between medium and complex gestures was found (see Fig. 4).

Overall, this experimental result validated the assumption that finger modality with occluded visual information resulted in the higher errors of the AD, IPD, LD, and SHV in the trajectory-based task, compared to when using a finger-attached pen.

III. Proposed Method-Designing Auditory Feedback

From the prior work of Park *et al.* [28], the possibility of the enhancement of the hand gesture interaction through coupling the continuous auditory (frequency and amplitude) with depth information was evident. Results showed that the depth continuous auditory feedback was more effective than that of the discrete auditory feedback for hand-gesture circular-menu selection. The circular-menu interface was divided into several visual regions, in which the discrete and continuous auditory feedbacks were coupled. When user's hand gesture crossed the boundary of each visual region, the auditory feedback was triggered to inform users. Based on this observation, we designed discrete beep, static, and gradual continuous auditory feedback for trajectory-based bare finger gestures to compensate for the occluded visual feedback in touch user interfaces. The relation between the gesture pattern and auditory feedback is established and described as follows.

Beep: Beep sound is mainly used for informing a transition event [13], [33] and a circular-menu selection [28], [29]. In this work, we therefore utilized the beep sound (discrete tone) to map with the endpoints and intersections of the template. The boundaries of endpoints and intersections are predefined according to the template gesture (the rectangle contained red dot represents the boundary of endpoints in Fig. 6 top). Each corner and intersection are coupled with the beep sound. When user's fingertip is tracing or crossing the area of corner or intersection of template, the discrete beep sound is given to notify users the drawing states.

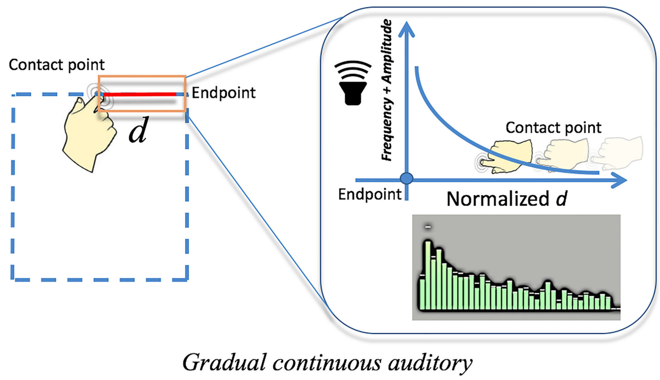


Fig. 7. The distance between contact point and predefined endpoint of the template is calculated and normalized, only applicable in this rectangle feedback area. The relationship between normalized distance and frequency and amplitude is illustrated.

Static: Distinguishing from discrete auditory feedback, a continuous auditory parameter, like the *frequency*, was selected for the representation of gesture tracing action. It is named after static continuous auditory feedback shown as follows:

$$f(t) = f_0 * e^t \tag{6}$$

where f_0 is the constant value, the frequency *f(t)* is used to represent the drawing itself. It remains at a steady level when the fingertip is in the predefined area of endpoints and intersections (see Fig. 6 bottom). The rectangle area is defined as the feedback area. When the user's fingertip traces the template in that area, the predefined static sound is triggered. Compared with the beep sound, more-continuous information regarding the drawing action itself can be given here. Users can easily recognize the continuous auditory feedback when tracing the predefined gestures on touchscreens.

Gradual: To provide more continuously informative feedback, a gradual-continuous auditory feedback is introduced to map the occluded visual feedback with *frequency* and *amplitude* cues. The relationship between the tracing point and the endpoint of the template gesture pattern using frequency and amplitude is shown as follows:

$$g_d(t) = W * e^{-d} * \cos[f(d) * t], \left(0 <= d <= \frac{l}{2}\right) \tag{7}$$

$$f(d) = f_0 * e^{\frac{l}{2} - d} \tag{8}$$

where (7) shows the relation between the generated gradual auditory information $g_d(t)$ and the spatial distance *d* in the predefined feedback area, where *d* is the normalized distance between the contact point and the endpoint of template, as shown in Fig. 7. The *d* can be calculated using the registered touch events in the developed Android application, and *l* represents the predefined length of gesture pattern. Equation (8) shows the relation between frequency *f (d)* and the *d*. The *W* and f_0 are constant value in the (7) and (8), for which two constants were utilized to ensure that the generated audio could be effective enough to be perceived by participants. The auditory feedback gradually changes depending on its distance from the drawn length. This is sufficient for the provision of the information about the distance between fingertip tracing point and the endpoint of template. If the tracing point is getting closer to the endpoint, the amplitude will gradually be increased (see Fig. 7). Compared with the static auditory feedback, the gradual continuous auditory feedback provides additional amplitude information. It is promising that this kind of continuous auditory cue delivers an incremental performance for compensating occluded visual feedback during a trajectory-based finger-gesture interaction.

IV. Experiment 2: Effects of Continuous Auditory Feedback on Trajectory-Based Finger Drawing

The effects of perceptual auditory feedback on the compensation of the visually occluded information for the improvement of the finger-modality accuracy in trajectory-based drawing tasks were designed. The individual and combination of auditory feedback served as the independent variable. The complexity of gestures served as the other independent variable. The same geometric features in the first experiment served as the dependent variables. The six feedback conditions for which visual feedback was available for each of them are shown as follows:

F1: No auditory feedback
F2: Beep auditory feedback (tone)
F3: Static auditory feedback (frequency)
F4: Gradual auditory feedback (frequency + amplitude)
F5: Beep + Static (tone + frequency)
F6: Beep + Gradual (tone + frequency + amplitude)

A. Participants and Apparatus

Fifteen students comprising four females and eleven males who were aged from 20 to 30 (average: 25.5, SD 2.4), with backgrounds in the field of computer science, participated in the experiment. Ten of participants were right-hand dominant, while the remaining participants were left-hand dominant. All of the participants have the normal hearing ability and can distinguish the five types of auditory feedback used in this experiment. They had also drawing experience using finger.

As with the first experiment, the same apparatus was used in this experiment, and an Android application was developed for predefined template gesture patterns with proposed different types of auditory feedback, as illustrated one of gesture patterns in Fig. 6. The touch event that registered the tracing position of the fingertip was used to calculate the normalized distance that finger drawn in that feedback area, as shown in Fig. 6 (bottom), then this normalized distance was used to calculate the amplitude and frequency with (7) and (8) (see Fig. 7). However, the width of feedback area became sensitive for triggering auditory feedback with fingertip, e.g., if the width is too large, the feedback area itself may influence the drawing performance. In contrast, participants find it hard to trace that area successfully. In this experiment, the width of feedback area was set as 0.5 cm.

B. Task and Procedure

This was the same as the first experiment. In the training stage, participants were asked to practice with the index finger of their dominant hand in the different feedback conditions for gesture patterns that were not used in the experiment. They were asked to report the type of the auditory feedback verbally after practicing. In the experimental stage, each participant was asked to draw the 12 gesture patterns randomly with the index finger of dominant hand by tracing the template under each feedback condition. Participants were required to draw them as accurately as possible under each of six feedback conditions

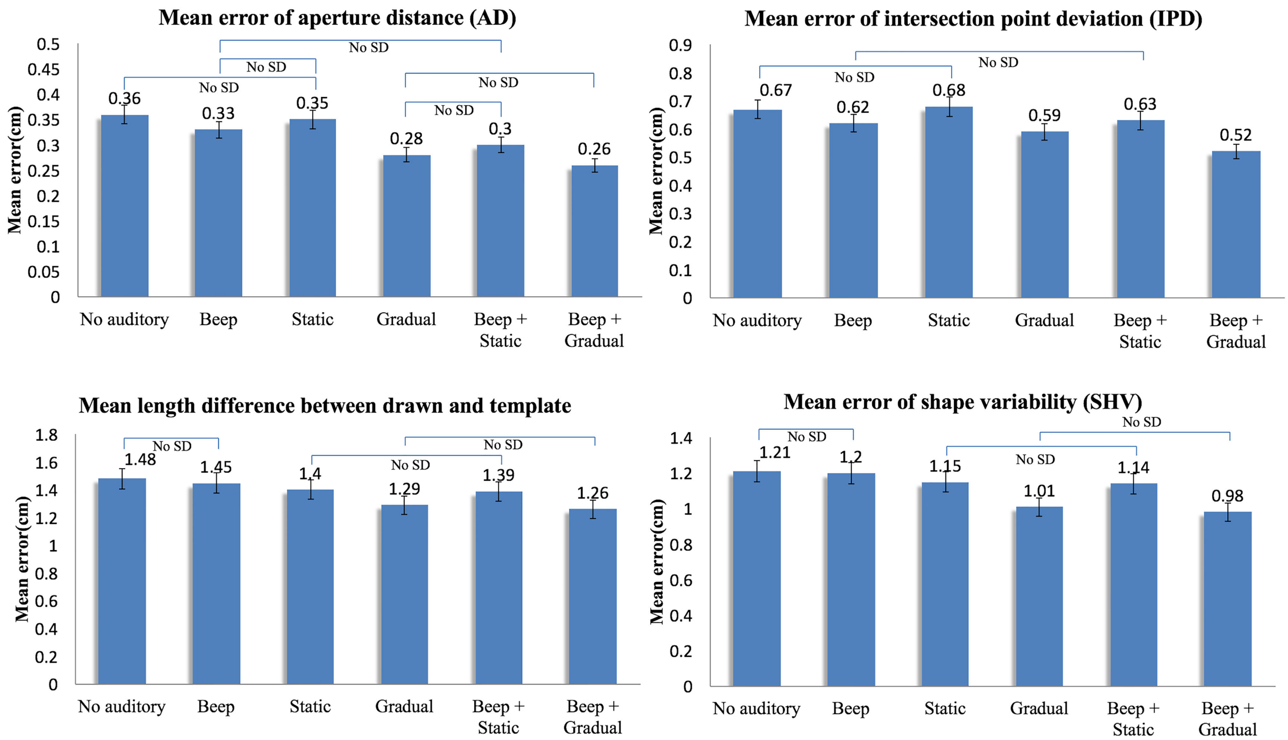


Fig. 8. The effects of different feedback conditions on finger modality in the trajectory-based tasks. Participants' performance in drawing gesture tasks on four geometric measures were collected. The results indicate that a combination of beep and gradual continuous auditory feedback contribute to higher accuracy of finger modality on all four measures. No significant differences (SD) for different types of auditory feedback on each measure marked with No SD, the detailed information of comparisons are shown in Table II. Error bar represents 95% confidence interval.

TABLE II
TUKEY'S HSD POSTHOC

Measures	F1 vs. F2	F1 vs. F3	F1 vs. F4	F1 vs. F5	F1 vs. F6	F2 vs. F3	F2 vs. F4	F2 vs. F5	F2 vs. F6	F3 vs. F4	F3 vs. F5	F3 vs. F6	F4 vs. F5	F4 vs. F6	F5 vs. F6
AD	*		*	*	*		*		*	*	*	*			*
IPD	*		*	*	**	*	*		**	*	*	*	*	*	*
LD		*	*	*	*	*	*	*	*	*		*	*		*
SHV		*	*	*	**	*	*	*	**	*		*	*		*

AD = Aperture Distance, *IPD* = Intersection Point Deviation, *LD* = Length Difference, *SHV* = Shape Variability.
* represents a significant difference at 0.05 level. ** represents a significant difference at 0.01 level. Blank space represents no significant difference.

Feedback conditions: F1 = no auditory, F2 = beep, F3 = static, F4 = gradual, F5 = beep + static, F6 = beep + gradual.

for four rounds. After each feedback condition, they had a 10-min break. Participants were asked to fill in the NASA-TLX questionnaire [37] on mental workload rating after finishing four rounds. The total number of drawing trials of this experiment was 4320 (15 subjects × 6 conditions × 12 gestures × 4 rounds = 4320).

C. Results and Analysis

The repeated measures ANOVA was adopted to measure the effects of the different types of auditory feedback. We further performed Tukey's HSD posthoc test for multiple comparisons between these conditions when significant differences were found on those measures.

Average gesture-completion time: The gesture-completion time with finger under combination of beep and gradual continuous auditory feedback was the shortest (3.14 s) across twelve gestures among six conditions. There was no significant difference of feedback conditions on the gesture-completion time ($F_{5,70} = 0.2, p = 0.17$). However, a significant difference of the gestural complexity on the gesture-completion time was found ($F_{2,28} = 62.4$, $p < 0.001$). Additionally, an interaction effect was found between the gestural complexity and the feedback conditions on the gesture-completion time ($F_{10,140} = 5.64, p < 0.01$). As expected, when the level of the gestural complexity increased, the gesture completion time became longer.

Aperture distance (AD): Fig. 8 (top left) shows the AD results under the 6 feedback conditions across the 11 closed gesture patterns. A significant difference for the feedback conditions on the AD was found: $F_{5,70} = 4.122, p < 0.05$. Posthoc test is shown in Table II. Conditions F2, F4, and F6 helped reduce the AD. In particular, the combination of beep and gradual-continuous auditory feedback (F6) achieved the highest performance of reducing AD error. Interestingly, the drawing performance with the static auditory feedback was far less effective than that with the use of the gradual auditory feedback, as participants re-

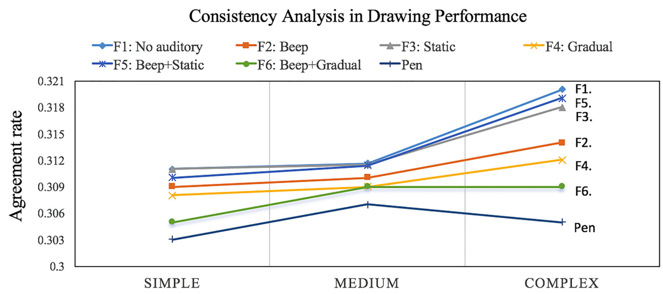


Fig. 9. Consistency analysis in drawing performance on each feedback condition across three groups of gestures. High agreement rate (RA) indicates that high consistency of making errors among four measures on that feedback condition. F1 reached the highest RA, while F6 remained at a similar level to the finger-attached pen modality.

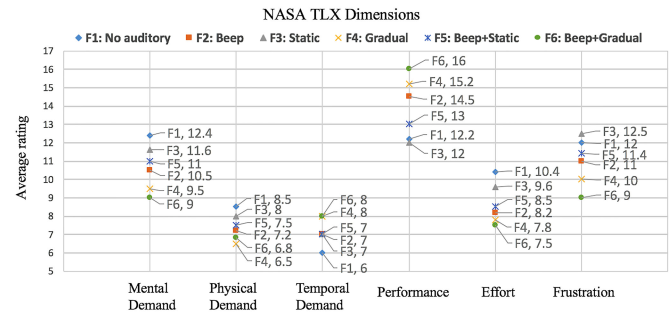


Fig. 10. Mean perceived mental workload ratings for the finger gestures under various feedback conditions across six NASA TLX dimensions, scores range from 1 to 20. The higher the score, the higher the perceived demand.

ported. However, no significant difference on the AD for the gesture complexity was found ($F_{2,28} = 0.32, p = 0.305$). There was no interaction effect between the gesture complexity and the feedback conditions on the AD ($F_{10,140} = 0.67, p = 0.405$).

Intersection-point deviation (IPD): The IPD performance across the feedback conditions is summarized in Fig. 8 (top right). A significant difference for the feedback conditions on the IPD was found ($F_{5,70} = 3.26$, $p < 0.05$). Posthoc test revealed that significant differences between all of condition pairs at $p < 0.05$ level, except for the comparisons of F1 versus F3 and F2 versus F5, as shown in Table II. Particularly, the combination of beep and gradual continuous auditory feedback greatly reduced the error of the IPD compared to no auditory feedback (F1 versus F6, $p < 0.01$). Participants found it difficult to use the static auditory feedback to help them draw more accurately at the intersection points within the small screen the smartphone provided. As with the AD, there was no significant difference for the gesture complexity on the accuracy of the IPD ($F_{2,28} = 0.48, p = 0.35$). An interaction effect between the gesture complexity and the feedback conditions on the IPD was not found ($F_{10,140} = 0.36, p = 0.405$).

Length difference (LD): Fig. 8 (Bottom-left) shows the LD across the feedback conditions. A significant difference for the feedback conditions on the LD was found ($F_{5,70} = 5.42$, $p < 0.05$). Posthoc test revealed significant differences between all the condition pairs at the $p < 0.05$ level, except for the comparisons of F1 versus F2, F3 versus F5, and F4 versus F6, as shown in Table II. It indicates that beep sound played a less important role in reducing the LD.

The significant difference for the gesture complexity on the LD was found ($F_{2,28} = 4.32, p < 0.05$). A significant difference between simple and complex gesture was found at $p < 0.01$. In contrast to the AD and IPD, an interaction effect was observed between the gesture complexity and the feedback conditions on the LD ($F_{10,140} = 2.23$, $p < 0.05$). Less LD errors were made when the participants traced the medium level of gestures under F4 condition, than under the F1, F2, F3, and F5 conditions. Condition F6 significantly reduced the LD errors compared to other conditions in the complex level of gestures. It revealed that the continuous auditory feedback was beneficial to the LD.

Shape variability (SHV): The summary of SHV value across the different feedback conditions is shown in Fig. 8 (bottom-right). There was a significant difference of the feedback conditions on the SHV ($F_{5,70} = 5.06$, $p < 0.05$). As with the LD, there was no effect of beep feedback on reducing the SHV value (see Table II). Both continuous auditory feedbacks helped reduce global features. In particular, the combination of beep and gradual continuous auditory feedback reached the lowest SHV value, which indicated that this feedback helped the participants producing the traces that were the closest to the predefined template. A significant difference for the gesture complexity on the SHV was found ($F_{2,28} = 5.05$, $p < 0.05$). Larger SHV value occurred in the complex gestures than that in the simple gestures ($p < 0.01$). There was an interaction effect between the gesture complexity and the feedback conditions ($F_{10,140} = 4.58$, $p < 0.05$). As expected, the F4- and F6-conditions reduced the SHV value in each of complexity level of gestures, compared to F1, F3, and F5.

Analysis of consistency in drawing performance: To measure user consistency in drawing gestures under each feedback condition, respectively, we adopted and refined the formula introduced by Anthony *et al.* [30] to calculate the agreement rate from both experiments. We utilized the drawing performance with pen in the first experiment as the ground truth. The calculation of the agreement rate for each feedback condition in the second experiment is done by equation (9). The $E = \{E_1, E_2, E_3, E_4\}$ contain the four geometric features. Then, the agreement rate in the feedback condition is defined in (9), where E_i represents the performance of *i* feature, *E* represents the sum performance of four features for that feedback condition

$$\mathrm{AR} = \sum_{i=1}^{4} \left(\frac{E_i}{E} \right)^2. \tag{9}$$

The higher agreement rate represents the higher consistency on making errors among four types of features. The consistency analysis in drawing performance under each feedback condition is shown in Fig. 9. The consistency trend of F6 condition is similar to visual feedback with finger-attached pen across three groups of gestures. The F2 and F4 conditions have a similar consistency trend across three groups of gestures. The other three conditions, F1, F3, and F5, remained at a similar consistent level. In particular, the more complex the gestures, the bigger the gap of performance across different feedback conditions. This consistency analysis reinforced the objective evaluation (see Fig. 8 and Table II).

Mental workload rating: The Wilcoxon signed rank test with Bonferroni correction analysis was employed to examine the mean difference between all condition pairs on each dimension of NASA-TLX [37]. Fig. 10 shows the mean of the participants' perceived workload for each dimension under each feedback condition. No significant differences between these feedback conditions on *physical* and *temporal* dimensions were found. As expected, a significant difference on *performance* was observed between all condition pairs at the $p < 0.003$ level, except for the comparison of F1 versus F5 conditions. There was a significant difference on *mental* demand between all condition pairs at the $p < 0.003$ level, except for the comparisons of F4 versus F6, F1 versus F3, F3 versus F5, and F2 versus F5 conditions. More *efforts* were required under both F1 and F3 conditions, compared to other feedback conditions F2, F4, F5, and F6 conditions ($p < 0.001$). The participants felt less *frustration* under F4 and F6 conditions, compared to F1, F2, F3, and F5 conditions ($p < 0.001$). To summarize, F6 condition obtained the best rating on these dimensions except *physical* and *temporal* dimensions.

V. Discussion

The first experiment indicated that visually occluded feedback led to the more errors in tracing the closed strokes and intersections of line segments using finger modality in the trajectory-based drawing tasks (see Fig. 4). The second experiment investigated the effects of different types of auditory feedback on trajectory-based finger gestures. The experimental results showed that the combination of beep and gradual continuous auditory performed as the best aid to help participants to draw trajectory-based finger gestures more precisely. In particular, using gradual continuous auditory feedback to improve both local and global geometric features was feasible, while beep sound improved local visual features only (see Fig. 8 and Table II).

A. Analysis of Auditory Parameters on Performance

Five types of auditory feedbacks were utilized in the second experiment. First, the three comparisons (F1 versus F2; F3 versus F5; and F4 versus F6) for effects of beep sound showed that tone increased the performance of on the AD and IPD (see Fig. 8 and Table II). Despite its advantage over the sole use of visual feedback, for example, the discrete beep feedback provided the participants with an alert or a notification as a status or an event transition. However, it was less effective for global geometric features, such as LD and SHV, as the movement requires accumulated knowledge regarding its trajectory and relation with the other already drawn gestures.

Participants reported that the beep sound was too transient, and it was therefore unable to provide a timely and correct reaction for the participants. This was consistent with Park *et al.* [28], though they investigated the mid-air finger gestures. The pre-emptive auditory feedback was proposed to provide user with a bit longer reaction time after the beep sound was released. In particular, three of participants failed to trigger the beep sound in the beginning round, due to unsuccessfully tracing the endpoints of the template. Accordingly, a further evaluation of pre-emptive auditory feedback on finger gestures will be conducted in the future work.

In an erroneous situation, the continuous auditory feedback provides additional information for the guiding of forthcoming movements toward the correct trajectory. Interestingly, the effect of static auditory feedback (frequency) on global geometric features did find significant difference compared with the effect of no auditory feedback (F1). Comparison of conditions F2 versus F5 indicated that frequency did not assist local geometric features, but some of participants reported that it guides the drawing action itself, which may give emotional feedback during gesturing, such as pleasure or surprise [31]. This further exploration of the emotional response for which the auditory feedback on finger modality is used will be investigated in the future work.

In terms of the gradual auditory feedback (amplitude and frequency), most participants reported positive feedback, as the incremental continuous information is given when the fingertip is crossing the endpoint or the intersections. Comparisons of F3 versus F4, and F5 versus F6 showed the effects of amplitude on each geometric features, and it helped the overall accuracy of drawing performance (see Figs. 8, 9 and Table II).

B. Modeling Trajectory-Based Finger Gestures Performance

A significant difference for the discrete and continuous auditory feedback on movement time was not found, while it was found on the accuracy of trajectory-based drawing (see Fig. 8). This was consistent with [33], which investigated the effects of multimodal feedback on steering law with stylus. It showed that the error feedback (discrete sound, tactile feedback) improved the accuracy of steering task other than time performance. However, the experimental design and the drawing implements were different (e.g., stylus with discrete sound and tactile error feedback [33]). In future work, the effects of tactile or haptic feedback [40] on trajectory-based finger gestures will be investigated.

One more concern on designing the continuous auditory feedback was the width and amplitude of the auditory feedback area. The interaction performance may differ with varying configuration. Prior research works have been explored on the factors of steering law [32], for example, the narrowing and widening tunnels; it derived and verified the relationship between navigation direction and index of difficulty. It was predefined in our experiment. In future work, we will investigate the effects of various width and amplitude of feedback area on drawing performances.

C. Implications of Continuous Auditory Feedback

In real environments, the endpoint could be expected when the finger or pen's pressure to the screen is lower than some value for a given period (expecting to lift up the finger or pen). In a real situation, it could be implemented that no additional sound feedback would be given after the expected endpoint is passed. However, the accuracy of endpoint could be enhanced by real-time recognition of writing patterns/symbols along with pressure changes, and it is not the overall scope of this paper.

In this paper, scenarios that illustrate the potential of continuous auditory feedback include touchscreen keyboards, menu interfaces, and driving situations.

On touchscreen keyboard, gesture typing is a common and efficient way of inputting English letters, such as *ShapeWriter*, *Swiftkey,* and Google Keyboard. Compared to touch typing, gesturing typing offers several advantages, as Smith *et al.* [38, p. 3365] stated,

> it supports a gradual and seamless transition from visually guided tracing to recall-based gesturing, allows users to approximate words with gestures rather than tapping each key exactly, but leaves a major problem compared to common typing: highly ambiguous word gestures.

In particular, the issue of "fat finger" leads to more errors when drawing with finger on such small touchscreens. Therefore, the designed continuous auditory feedback could give more information to guide tracing on small touchscreens. For example, the gradual continuous auditory feedback can be mapped with the layout of keys, specifically, the spatial location of each key on the keyboard layout can be mapped with frequency and amplitude. When the user's fingertip is tracing on such layout, the corresponding auditory feedback will be given to inform the position of fingertip. Therefore, it can reduce the space of layout and increase the performance.

In the circular or hierarchy menus of touch user interfaces, some prior works have shown the advantages of sound feedback [29]. For example, the well-known "eyes-free" menu selection technique using reactive audio feedback [29], and Brewster [13] also found that the additional sound to buttons for mobile interactions was beneficial. It revealed advantages of using discrete sound for menu selections in touch user interfaces. However, it is potential for the use of gradual continuous auditory feedback for improving circular or hierarchy menu selection on small touchscreens.

In driving or navigating situations [34], [41], when backing up a car into a parking lot, the distance between obstacle and car can be detected in real time. Even though the discrete beep sound is already known to provide warning for drivers, it is not enough to give timely feedback sometimes. In particular, the distance is too closer to allow the driver to react within the time it takes for the next beep to play. In such a case, a gradual continuous auditory could provide a more persistent information. The continuous sound is changing gradually according to the distance between the car and the obstacle.

VI. Conclusion

This paper first investigated the effects of concurrent visual feedback between the finger and a finger-attached pen in trajectory-based drawing tasks. The results from the experiments indicated that visually occluded feedback contributed to a low accuracy of finger modality in drawing geometric features, such as closed gestures, intersections of gesture patterns, and gesture's global difference (LD and SHV). Then, different types of auditory feedback (discrete beep, static, gradual) that represent different perceptual cues are proposed to improve the drawing performance of the trajectory-based finger gestures. The gradual continuous auditory feedback performed better on the participants' drawing performances than both the static- and beep-sound. Particularly, the combination of beep (tone) and gradual continuous auditory (amplitude + frequency) feedback that provides incremental and stated information helped users perform the best drawing performance among all the tested feedback conditions. The potential uses of the findings were discussed for touchscreen keyboards, circular menu design, and driving situations.

References

[1] H. S. Vitense *et al.*, "Multimodel feedback: An assessment of performance and mental workload," *Ergonomics*, vol. 46, pp. 68–87, 2003.

[2] K. A. Siek *et al.*, "Fat finger worries: How older and younger users physically interact with PDAs," in *Proc. IFIP Conf. Human-Computer Interaction*, Rome, Italy, 2005, pp. 267–280.

[3] H. Tu *et al.*, "A comparative evaluation of finger and pen stroke gestures," in *Proc. SIGCHI Conf. Human Factors Comput. Syst.,* Austin, TX, USA, 2012, pp. 1287–1296.

[4] A. Cockburn *et al.*, "Understanding performance in touch selections: Tap, drag, and radial pointing drag with finger, stylus, and mouse," *Int. J. Human-Comput. Stud.*, vol. 70, no. 3, pp. 218–233, 2012.

[5] J. Zhao *et al.*, "A model of scrolling on touch-sensitive displays," *Int. J. Human-Comput. Stud.*, vol. 72, no. 12, pp. 805–821, 2014.

[6] P. Albinsson *et al.*, "High precision touch screen interaction," in *Proc. SIGCHI Conf. Human Factors Comput. Syst.*, Ft. Lauderdale, FL, USA, 2003, pp. 105–112.

[7] H. Benko *et al.*, "Precise selection techniques for multi-touch screens," in *Proc. SIGCHI Conf. Human Factors Comput. Syst.*, Montreal, Canada, 2006, pp. 1263–1272.

[8] D. Vogel *et al.*, "Shift: A technique for operating pen-based interfaces using touch," in *Proc. SIGCHI Conf. Human Factors Comput. Syst.*, San Jose, CA, USA, 2007, pp. 657–666.

[9] C. Holz *et al.*, "The generalized perceived input point model and how to double touch accuracy by extracting fingerprints," in *Proc. SIGCHI Conf. Human Factors Comput. Syst.*, Atlanta, GA, USA, 2010, pp. 581–590.

[10] B. Gao *et al.*, "Use of sound to provide occluded visual information in touch gestural interfaces," in *Proc. 33rd Annu. ACM Conf. Extended Abstr. Human Factors Comput. Syst.*, Seoul, South Korea, 2015, pp. 1277–1282.

[11] H. Xia *et al.*, "NanoStylus: Enhancing input on ultra-small displays with a finger-mounted stylus," in *Proc. 28th Annu. ACM Symp. User Interface Softw. Technol.*, Charlotte, NC, USA, 2015, pp. 447–456.

[12] T. H. Andersen *et al.*, "Writing with music: Exploring the use of auditory feedback in gesture interfaces," *ACM Trans. Appl. Perception*, vol. 7, no. 17, 2010, Art no. 17.

[13] S. Brewster, "Overcoming the lack of screen space on mobile computers," *Pers. Ubiquitous Comput.*, vol. 6, no. 3, pp. 188–205, 2002.

[14] E. Thoret *et al.*, "From sound to shape: Auditory perception of drawing movements," *J. Exp. Psychol. Human Perception Perform.*, vol. 40, pp. 983–994, 2014.

[15] U. Oh *et al.*, "Audio-based feedback techniques for teaching touchscreen gestures," *ACM Trans. Access. Comput.*, vol. 7, 2015, Art no. 9.

[16] X. Bi *et al.*, "FFitts law: Modeling finger touch with fitts' law," in *Proc. SIGCHI Conf. Human Factors Comput. Syst.*, Paris, France, 2013, pp. 1363–1372.

[17] P. M. Fitts, "The information capacity of the human motor system in controlling the amplitude of movement," *J. Exp. Psychol.*, vol. 47, no. 6, pp. 262–269, 1954.

[18] H. Song *et al.*, "Grips and gestures on a multi-touch pen," in *Proc. SIGCHI Conf. Human Factors Comput. Syst.*, Vancouver, BC, Canada, 2011, pp. 1323–1332.

[19] L. Stearns *et al.*, "Evaluating haptic and auditory directional guidance to assist blind people in reading printed text using finger-mounted cameras," *ACM Trans. Access. Comput.*, vol. 9, no. 1, 2016, Art no. 1.

[20] X. Cao *et al.*, "Modeling human performance of pen stroke gestures," in *Proc. SIGCHI Conf. Human Factors Comput. Syst.*, San Jose, CA, USA, 2007, pp. 1495–1504.

[21] R. Pastel, "Measuring the difficulty of steering through corners," in *Proc. SIGCHI Conf. Human Factors Comput. Syst.*, Montreal, Canada, 2006, pp. 1087–1096.

[22] R. S. Kattinakere *et al.*, "Modeling steering within above-the-surface interaction layers," in *Proc. SIGCHI Conf. Human Factors Comput. Syst.*, San Jose, CA, USA, 2007, pp. 317–326.

[23] A. Cockburn *et al.*, "Multimodal feedback for the acquisition of small targets," *Ergonomics*, vol. 48, no. 9, pp. 1129–1150, 2005.

[24] M. Annett *et al.*, "Hands, hover, and nib: Understanding stylus accuracy on tablets," in *Proc. 41st Graph. Interface Conf.*, 2015, pp. 203–210.

[25] Pen stylus, 2012. [Online]. Available: http://www.adonit.net/jot/. Accessed on: Mar. 25, 2016.

[26] M. Smyth *et al.*, "Function of vision in the control of handwriting," *Acta Psychologica*, vol. 65, pp. 47–64, 1987.

[27] R. D. Vatavu *et al.*, "Relative accuracy measures for stroke gestures," in *Proc. 15th ACM Int. Conf. Multimodal Interact.*, Sydney, Australia, 2013, pp. 279–286.

[28] Y. Park, J. Kim, and K. Lee, "Effects of auditory feedback on menu selection in hand-gesture interfaces," *IEEE Multimedia*, vol. 22, no. 1, pp. 32–40, Jan.–Mar. 2015.

[29] S. Zhao *et al.*, "Earpod: Eyes-free menu selection with touch input and reactive audio feedback," in *Proc. SIGCHI Conf. Human Factors Comput. Syst.*, San Jose, CA, USA, 2007, pp. 1395–1404.

[30] L. Anthony *et al.*, "Understanding the consistency of users' pen and finger stroke gesture articulation," in *Proc. Graph. Interface*, Regina, Sascatchewan, Canada, 2013, pp. 87–94.

[31] T. Jimenez *et al.*, "Using sound in multi-touch interfaces to change materiality and touch behavior," in *Proc. 8th Nordic Conf. Human-Comput. Interaction: Fun, Fast, Foundational*, Helsinki, Finland, 2014, pp. 199–202.

[32] S. Yamanaka *et al.*, "Modeling the steering time difference between narrowing and widening tunnels," in *Proc. SIGCHI Conf. Human Factors Comput. Syst.*, San Jose, CA, USA, 2016, pp. 1846–1856.

[33] M. Sun *et al.*, "Effects of multimodal error feedback on human performance in steering tasks," *J. Inf. Process.*, vol. 18, pp. 284–292, 2010.

[34] M. J. Jensen, A. M. Tolbert, J. R. Wagner, F. S. Switzer, and J. W. Finn, "A customizable automotive steering system with a haptic feedback control strategy for obstacle avoidance notification," *IEEE Trans. Veh. Technol.*, vol. 60, no. 9, pp. 4208–4216, Nov. 2011.

[35] E. O. Boyer *et al.*, "Investigating three types of continuous auditory feedback in visuo-manual tracking," *Exp. Brain Res.*, vol. 235, pp. 691–701, 2017.

[36] G. Parseihian, C. Gondre, M. Aramaki, S. Ystad, and R. Kronland-Martinet, "Comparison and evaluation of sonification strategies for guidance tasks," *IEEE Trans. Multimedia*, vol. 18, no. 4, pp. 674–686, Apr. 2016.

[37] S. G. Hart *et al.*, "Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research," in *Human Mental Workload*, P. A. Hancock and N. Meshkati, Eds. Amsterdam, The Netherlands: Elsevier Sci., 1988, pp. 139–183.

[38] B. A. Smith *et al.*, "Optimizing touchscreen keyboards for gesture typing," in *Proc. SIGCHI Conf. Human Factors Comput. Syst.*, Seoul, South Korea, 2015, pp. 3365–3374.

[39] M. Aramaki, M. Besson, R. Kronland-Martinet, and S. Ystad, "Controlling the perceived material in an impact sound synthesizer," *IEEE Trans. Audio, Speech, Language Process.*, vol. 19, no. 2, pp. 301–314, Feb. 2011.

[40] A. Crossan *et al.*, "Multimodal trajectory playback for teaching shape information and trajectories to visually impaired computer users," *ACM Trans. Access. Comput.*, vol. 1, no. 2, 2008, Art no. 12.

[41] S. Joseph *et al.*, "Be aware of the world: Toward using social media to support the blind with navigation," *IEEE Trans. Human–Mach. Syst.*, vol. 45, no. 3, pp. 399–405, 2015.

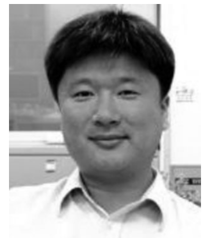
[42] R. Sigrist *et al.*, "Augmented visual, auditory, haptic, and multimodal feedback in motor learning: A review," *Psychonomic Bull. Rev.*, vol. 20, no. 1, pp. 21–53, 2013.

[43] R. Sigrist *et al.*, "Sonification and haptic feedback in addition to visual feedback enhances complex motor task learning," *Exp. Brain Res.*, vol. 233, no. 3, pp. 909–925, 2014.

[44] E. Freeman *et al.*, "Multimodal feedback in HCI: Haptics, non-speech audio and their applications," in *The Handbook of Multimodal-Multisensor Interfaces.* San Rafael, CA, USA: ACM and Morgan & Claypool, pp. 277–317, 2017.

BoYu Gao (S'17) received the Ph.D. degree in internet and multimedial engineering from Konkuk University, Seoul, South Korea, in February 2018.

He is currently a Postdoctoral Researcher with the Department of Software, Konkuk University. His research interests include designing multimodal feedback mechanism and evoking believable responses in virtual environments.



HyungSeok Kim (M'01) received the Ph.D. degree in computer science from Korea Advanced Institute of Science and Technology, Daedeok, South Korea, in February 2003.

Since 2006, he has been a Professor with the Department of Software, Konkuk University, Seoul, South Korea. During 2003 to 2006, he was a Senior Researcher with the MIRALab, University of Geneva. His research interests include real-time interaction in virtual environments and multimodal interaction mechanisms, shape modeling for real-time rendering, and evoking believable experiences in VEs.

Hasup Lee received the Ph.D. degree in computer science from Korea Advanced Institute of Science and Technology, Daedeok, South Korea, in 2007.

He is currently a Research Fellow with the VR Lab, Konkuk University, Seoul, South Korea. He was a Senior Assistant Professor with the Visual Simulation Lab, Keio University. His research interests include VR, computer graphics, and computer vision.

Jooyoung Lee received the Ph.D. degree in internet and multimedial engineering from Konkuk University, Seoul, South Korea, in February 2018.

He is currently a Postdoctoral Researcher with the Department of Software, Konkuk University. His research interests include 3-D reconstruction from multistereo and AR technology.

Jee-In Kim (M'13) received the B.S. degree in computer engineering from Seoul National University, Seoul, South Korea, in 1982, the M.S. degree in computer science from Korea Advanced Institute of Science and Technology, Daedeok, South Korea, in 1984, and the Ph.D. degree in computer and information science from University of Pennsylvania, Philadelphia, PA, USA, in 1993.

He has been a Professor at Konkuk University, Seoul, Korea, since 1995. He is the Founding Director of Digital Contents Research Center, Konkuk University. His research interests include HCI, VR, natural UIs, accessibility, and data visualization.