Use of Sound to Provide Occluded Visual Information in Touch Gestural Interface

BoYu Gao
Konkuk University
120, Neung-dongro
Seoul, South Korea
yuye@konkuk.ac.kr

HyungSeok Kim
Konkuk University
120, Neung-dongro
Seoul, South Korea
hyuskim@konkuk.ac.kr

Hasup Lee
Konkuk University
120, Neung-dongro
Seoul, South Korea
hasups@gmail.com

Jooyoung Lee
Konkuk University
120, Neung-dongro
Seoul, South Korea
ljy1201@konkuk.ac.kr

Jee-In Kim
Konkuk University
120, Neung-dongro
Seoul, South Korea
jnkm@konkuk.ac.kr

CHI'15 Extended Abstracts, Apr 18-23, 2015, Seoul, Republic of Korea
ACM 978-1-4503-3146-3/15/04
http://dx.doi.org/10.1145/2702613.2732817

Abstract
Direct touch gestures are getting popular as an input modality for mobile and tabletop interaction. However, the finger touch interface is considered as not accurate compared with pen-based interface. One of the main reasons is that the visual feedback of the finger touch is occluded because of the size of fingertip. It has made difficult for perceiving and correcting errors. We propose to utilize another modality to provide information on occluded area. Spatial information on visual channel is transformed to temporal and frequency information on another modality. We use sound modality to illustrate the proposed trans-modality. Results show that performance with additional modality is better for drawing where the visual information is important than only with the visual feedback.

Author Keywords
Visual occlusion, sound feedback; finger gestures; touch gestural interface.

ACM Classification Keywords
H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

Introduction

As platform for interaction moves to mobile and other pervasive computing forms, like wearable computing and smart environment, direct touch interaction has been the focus of numerous research projects [1,2,3,4,5] and appears to offer users more natural and intuitive experience. Recently, it becomes the most preferred metaphor for everyday HCI applications. However, direct touch gestural interaction on the small screen of smart devices is often very visually demanding to use. It is difficult to accurately point at objects that are smaller than finger on such small screen. To solve this issue, clever techniques have been developed to alleviate the issue [4,5]. For example, Albinsson [4] proposed two techniques Cross-Keys and Precision-Handle to allow pixel level pointing in a fast and efficient manner. And Hrvoje [5] proposed a technique Dual Finger Selection to facilitate pixel accurate targeting by adjusting the control display ratio with a secondary finger. However, the limitation of those approaches is lack of continuous and onsite visual feedback for finger strokes. The occlusion is considered to be one of important factors to influence continuous visual feedback [1]. Researchers have suggested that occlusion likely contributes errors, increases fatigue, forces inefficient movements and impedes performance [1,2,3]. Specifically, Tu [9] did comparative evaluation of stroke gestures with finger and pen, and found that it is hard to draw details with finger.

Therefore, designing less visually dependent user interface through more effective use of other modalities becomes more necessary than ever. For example, Anderson [6] investigated the use of auditory feedback in pen gesture interface and showed that gaining performance advantage with auditory feedback is possible. Sarah [7] explored the auditory cues to support motion gesture interaction on mobile devices. They proposed two techniques that use audio features to provide spatial representation of desired gesture and feedback on the system's interpretation of user input. Ana [8] explored the sound based interaction to alter the perceived materiality of a surface being touched and to shape user's touch behavior. Those researches show promising results on utilizing sound modality as a secondary interaction channel. It is still an open issue that how the secondary channel can provide missing information on the primary channels such as occluded shapes.

Given those open questions, we perform an investigation on touch gestural interaction to assist the accuracy of finger interaction. Specifically, drawing details with finger is supposed to be hard to reach user's expectation due to occluded visual information. Compared with existing approaches [4, 5], we put focus on devising method to provide additional information to compensate the occluded visual information. The novelty of our idea is to use temporal sound cues to provide occluded spatial visual information. For example, the visual spatial distance can be transferred into temporal sound volume intensity accordingly. The detail of proposed method shows how the sound feedback works in following section.

Proposed method

The goal of this work is to explore the sound cues to provide occluded visual information to assist accuracy in touch gestural interface. The key issues are on the kind of sound cues which can be used to provide occluded spatial visual information.

Types	Features
Local shape information	Aperture distance between start point and end point of closed gesture
Global shape information	Deviation of intersection point

Table. 1. Occluded visual information

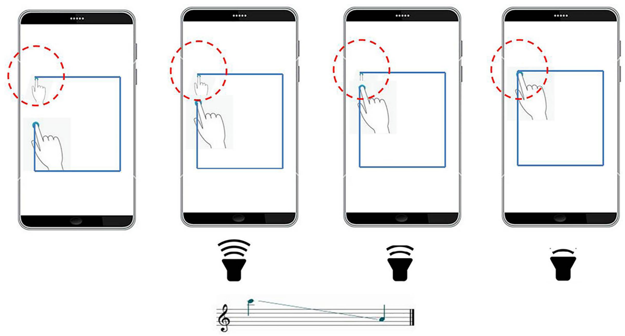

Fig 1.a. Corresponding sound information is provided for aperture distance between starting point and end point of closed gesture.

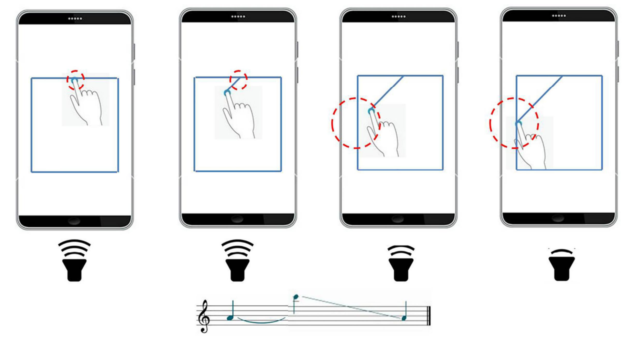

Fig 1.b. Corresponding sound information is provided for deviation of intersection point.

Observations

The comparative study of stroke gesture with finger or pen has shown the similarities and differences [9]. Based on the findings [1,2,9], we summarize these following observations:

- It is highly possible to make errors for selecting or pointing the object which is smaller than fingertip directly with user's finger.
- It is hard to draw the details with touch gesture interaction. The level of accuracy is relatively low.
- Direct touch interaction involves high visual attention in real environment.
- Sound modality can provide continuous and informative feedback.

Analysis of occluded visual features

Due to the contact area and size of fingertip, the continuous visual feedback is not always feasible during touch gesture interaction. Tu [9] listed some visual features as important cues, such as aperture distance between two endpoints of closed stroke and deviation of intersection point. Those features in Table.1 are assumed to be important factors to influence the accuracy during touch gesturing. In this work, we use the spatial distance between tracing point and starting point as an exemplar feature as shown in Fig 1.a. In addition, we also investigate the deviation of intersection point between two strokes in Fig 1.b.

Sound cues to provide occluded visual features

We develop a prototype to provide sound temporal information to assist the accuracy of finger strokes according to analysis of occluded visual features. Specifically, the two types of sound feedback are provided when you contact the area which defined to meet the following situations:

- In the Fig 1.a, the red dotted circle represents the triggered area. The sound feedback will be given when the finger contact point is in this circle. The distance between tracing point and starting point is formulated

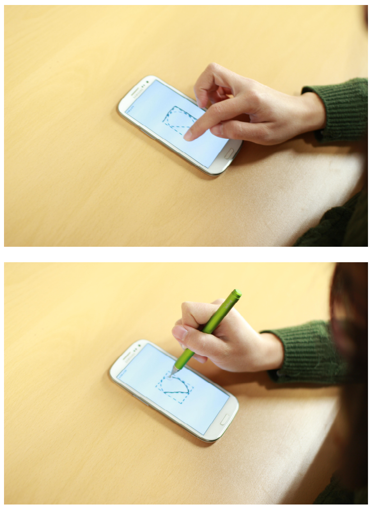

Fig 2. Finger & Pen drawing

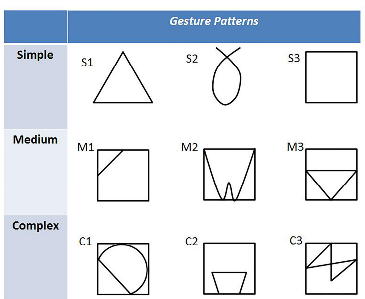

Fig 3. Gesture pattern

into the intensity of volume as shown in Fig 1.a. If the tracing point is getting closer to the starting point, the intensity of volume will be decreased accordingly.

- The selection of smaller soft button is hard for the user whose fingertip is fat. In the first figure of Fig 1.b, we define the small circle as the intersection area to notice user with sound feedback. If the user contacts this area, the beep sound is given.
- After selecting the intersection point, user starts moving the touch stroke forward as shown in Fig 1.b, if the tracing point is getting closer to the predefined circled area, the temporal sound information will be provided accordingly.

Pilot study - Evaluation of occluded visual information

The goal of this study is to validate the occluded visual information during touch gesturing. We assume that the accuracy of stroke gesture is decreased accordingly due to lack of visual feedback.

Conditions

In order to eliminate the influenced factors as many as possible, the participants are required to follow the guideline to draw with finger and pen as accurately as possible like Fig.3 shown. On the other hand, the angle between pen and touch screen should be kept around 45 degree like the Fig.2 shown, in order to show the visual feedback for pen drawing as much as possible.

Apparatus and participants

The study was conducted on a Galaxy S3 mobile phone with 4.8-inches, 720×1280 pixels, which can support finger and pen drawn gestures. We recruited fifteen volunteers, five females and ten males, from 20 to 50 years of age, with background in the field of art design, computer science, and international trade, participated in the experiment. Four of them had prior experience using both pen stylus and finger operation. Eight of them had prior experience either pen stylus or finger operation. Three of them had no experience using neither finger nor pen stylus.

Procedure

In order to eliminate the learning effects, we firstly asked participants to get familiar with finger and pen drawing. 20 minutes training stage is given for each participant. Then, they were asked to draw each gesture one by one for four blocks. After finishing one block, ten minutes break and a questionnaire about subjective evaluation were provided.

Discussion and Results

The result shows the difference between finger and pen drawing in terms of gesture complexity. Specifically, we observed the aperture distance and intersection

Type	Simple	Medium	Complex
Finger(Aperture)	0.40cm	0.65cm	0.6cm
Pen(Aperture)	0.26cm	0.42cm	0.36cm
Finger(Intersection)	0.36cm	1.02cm	1.07cm
Pen(Intersection)	0.21cm	0.56cm	0.65cm

Table 2. Comparison results of finger and pen in terms of mean error of aperture distance and intersection point deviation.

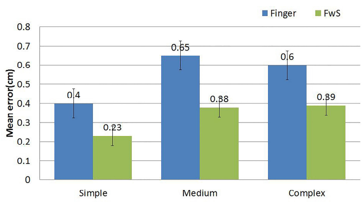

Fig 4. Aperture distance

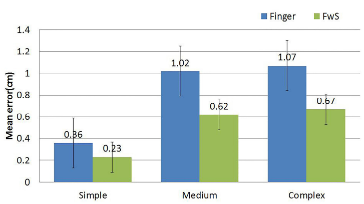

Fig 5. Deviation of Intersection point

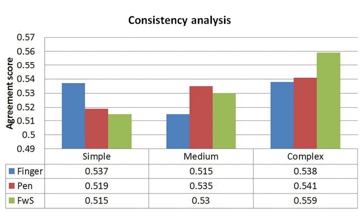

Fig 6. Consistency analysis

deviation between finger and pen for each level of complexity which shown in the Table.2.

We found there is no big difference in simple gesture for both features. While for medium and complex gesture, the mean error of intersection point deviation is almost two times of aperture distance as shown in Table.2. Interestingly, the result shows the mean error of medium gesture in terms of aperture distance is higher than complex gesture (Finger: 0.65cm vs. 0.6cm; Pen: 0.42cm vs. 0.36cm in Table.2).

Pilot study - Evaluation of sound feedback

The goal of this study is to evaluate performance of sound feedback to provide occluded visual information in touch gestural interface. We developed the prototype to support sound feedback with predefined touch gesture in Java using Android SDK.

Conditions and Procedure

The apparatus and participants are the same as previous evaluation. Additionally, the procedure is also the same as previous one except the implement modality (finger with sound feedback).

Discussion and Results

In order to evaluate the performance of sound feedback, we propose that two aspects are considered to evaluate. The one is the accuracy of stroke gesture with sound providing. The other one is consistency on user performance with each modality.

The result of the study shows the accuracy of finger with sound is better than finger in terms of aperture distance and deviation of intersection point as the Fig.4 & 5 & Table.2 shown. Interestingly, the mean error of aperture distance for simple and medium gesture with sound compensation is even better than pen (Simple: Finger with Sound vs. Pen is 0.23cm vs. 0.26cm; Medium: Finger with Sound vs. Pen is 0.38cm vs. 0.40cm). However, in terms of complex gesture, the performance of pen is the best for both aperture and intersection point drawing. Compared with the two features in terms of sound compensation, the aperture drawing is performed better than intersection point drawing for simple and medium gestures.

The other aspect is the consistency on user performance in terms of each modality in different level of gestures. We employed the formula introduced by Anthony [10] for calculating an agreement score for each modality under different levels of complexity. The result is shown in Fig.6. The higher score represents the high consistency with that modality for two features. For example, participants prefer the simplistic and convenient of finger drawing. As the result shows, the finger performs well for both aperture and intersection features in simple gestures even though the accuracy of finger is lowest for both features. Although the lowest agreement on finger with sound for simple gesture, in other words, some of users prefer it, while others are not. Interestingly, the accuracy of finger with sound is relatively higher than pen (Finger with Sound vs. Pen is 0.23cm vs. 0.26cm). In terms of medium gesture, there is no big difference. The agreement score of pen is the highest also for the accuracy. For the complex gesture, the highest agreement score is finger with sound. However, the highest accuracy of both features is pen drawing. The implication of result shows the accuracy is not directly connected with user preference.

Conclusion and Future work

The focus of this work is to explore sound cues for occluded visual information in touch gestural interface. We develop the prototype to prove our assumption through two pilot studies. Basically, the results show that proposed sound feedback has a good performance on both occluded visual features. Specifically, sound feedback performs better for drawing endpoints than intersection points in terms of simple and medium gestures. Additionally, the pilot study shows that pen is better for drawing endpoints of closed strokes in complex gesture.

Our work still has several limitations that point toward future work. Firstly, the designed gesture should be reconsidered due to the performance of aperture distance error (The mean error of Medium level is higher than complex). Secondly, the mapping between occluded visual information and temporal sound cues should be modeled according to the analyzed visual features. Additionally, more visual features should be explored, not only aperture distance between starting point and ending point but also deviation of intersection point. Thirdly, to consider haptic modality which can also provide continuous and informative cues in the loudly and noisy environment.

Acknowledgements

This research was supported by Next-Generation information Computing Development Program (No.2012M3C4A7032185) and Bio-Synergy Research Project (No.2013M3A9C4078140) through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT & Future Planning.

References

[1] Forlines, C., Wigdor, D., Shen, C., Balakrishnan, R. Direct-Touch vs. Mouse Input for Tabletop Displays. *In Proc. of CHI*, ACM Press(2007), 647- 656.

[2] Forlines, C. and Balakrishnan, R., Evaluating tactile feedback and direct vs. indirect stylus input in pointing and crossing selection tasks. *In Proc. of CHI,* ACM Press (2008), 1563-1572.

[3] Vogel, D. and Balakrishnan, R., Occlusion-aware interfaces. *In Proc. of CHI,* ACM Press (2010), 263-272.

[4] Albinsson, P. and Zhai, S. High precision touch screen interaction. *In Proc. of CHI,* ACM Press (2003), 105-112.

[5] Benko, H., Wilson, A., and Baudisch, P. Precise selection techniques for multi-touch screens. *In Proc. of CHI,* ACM Press (2006), 1263-1272.

[6] Andersen, T.H. and Zhai, S. “Writing with music”: exploring the use of auditory feedback in gesture interfaces. ACM *TAP 7*, 3(2010), 1-24.

[7] Sarah, M.S., Ruiz, J., Using Audio Cues to Support Motion Gesture Interaction on Mobile Devices, *In Proc. of CHI,* ACM Press (2014), 1621-1626.

[8] Tajadura-Jimenez, A., Liu, B., Bianchi-Berthouze, N., and Bevilacqua, F. Using sound in multi-touch interfaces to change materiality and touch behavior, *In Proc. of NordiCHI*, ACM Press (2014), 199-202.

[9] Tu,H., Ren,X., Zhai,S. A comparative evaluation of finger and pen stroke gestures, In *Proc. of CHI*, ACM Press (2012), 1287-1296.

[10] Anthony,L., Vatavu, R.D., and Wobbrock, J.O, Understanding the consistency of users’ pen and finger stroke gesture articulation, In Proc of the Graphical Interface 2013, Canadian Information Processing society, 2013, 87-94.