國立臺灣大學電機資訊學院資訊網路與多媒體研究所

碩士論文

Graduate Institute of Networking and Multimedia
College of Electrical Engineering & Computer Science
National Taiwan University

Master Thesis

智慧型眼鏡在公眾場合中的使用者定義遊戲操作
User-Defined Game Input for Smart Glasses in Public Space

王瀚宇
Han-Yu Wang

指導教授：陳彥仰博士
Advisor: Mike Y. Chen, Ph.D.

中華民國 104 年 6 月
June, 2015

國立臺灣大學碩士學位論文
口試委員會審定書

智慧型眼鏡在公眾場合中的使用者定義遊戲操作
User-Defined Game Input
for Smart Glasses in Public Space

本論文係王瀚宇君 (R02944002) 在國立臺灣大學資訊網路與多媒體研究所完成之碩士學位論文，於民國 104 年 6 月 2 日承下列考試委員審查通過及口試及格，特此證明

口試委員：

所　長：

誌謝

感謝實驗室的夥伴們對於我的幫助，尤其是同組的夥伴們，我們互相學習、互相加油打氣，我們一起努力的日子是我最難忘的回憶。也感謝我的指導教授陳彥仰老師，給予了我許多的建議，且十分支持我的研究。感謝所有在過程中給予我幫忙的朋友，你們的每一分付出，都成就了這篇論文的每字每句，謝謝你們。

摘要

智慧型眼鏡（如：Google Glasses），具備有

Abstract

Smart glasses, such as Google Glass, provide always-available displays not offered by console and mobile gaming devices, and could potentially offer a pervasive gaming experience. However, research on input for games on smart glasses has been constrained by the available sensors to date. To help inform design directions, this paper explores user-defined game input for smart glasses beyond the capabilities of current sensors, and focuses on the interaction in public settings. We conducted a user-defined input study with 24 participants, each performing 17 common game control tasks using 3 classes of interaction and 2 form factors of smart glasses, for a total of 2448 trials. Results show that users significantly preferred non-touch and non-handheld interaction over using handheld input devices, such as in-air gestures. Also, for touch input without handheld devices, users preferred interacting with their palms over wearable devices (51% vs 20%). In addition, users preferred interactions that are less noticeable due to concerns with social acceptance, and preferred in-air gestures in front of the torso rather than in front of the face (63% vs 37%).

Contents

口試委員會審定書 **ii**

誌謝 **iii**

摘要 **iv**

Abstract **v**

1 Introduction **1**

2 Related Work **5**

2.1 Game Input 5

2.2 Mobile Input Technology 5

2.3 Gestures in HCI 6

2.4 User Elicitation Studies 6

Bibliography **8**

List of Figures

1.1 A study participant performing an in-air gesture to drag an object seen through the immersive smart glasses in a public coffee shop. 2

1.2 (A) Epson Moverio, (B) Google Glass. 2

List of Tables

Chapter 1

Introduction

Smart glasses provide always-available displays and offer the opportunity for instantly available information and pervasive gaming experiences. Compared to game consoles and mobile gaming devices, smart glasses do not have touchscreens and currently do not support handheld controllers specifically designed for gaming. Current smart glasses, such as Google Glass and the Epson Moverio, support input via voice, touchpads, cameras, gyroscopes, accelerometers, and GPS. Games designed specifically for Google Glass [17] utilize these sensors as game control. For example, “Clay Shooter” utilizes the user’s voice to trigger a shotgun, and “Shape Splitter” detects in-air gestures via the built-in cameras. For the Epson Moverio glasses, wired trackpads are used as handheld inputs.

To better inform the interaction design of games for smart glasses, we aimed to explore the design space without being constrained by the capabilities of current sensors. We used the guessability study methodology [26], and presented the *effects* of game controls to the participants in a real-world, public environment. We then elicited what the participants felt was the most appropriate *causes* to invoke the corresponding effects.

For input tasks, we analyzed 90 popular games to identify the game controls used by more than one game, which resulted in a set of 17 tasks. We also explored the form factors of smart glasses displays, and included both types in the study: 1) *immersive*, with display content spanning the user’s field of view (e.g. Epson and Sony’s smart glasses), and 2) *off-to-the-side*, with display content in the corners of the user’s field of view (e.g. Google Glass).

Figure 1.1: A study participant performing an in-air gesture to drag an object seen through the immersive smart glasses in a public coffee shop.

In order to compare different types of interaction while keeping the experiment tractable, we grouped the different types of input into the following 3 classes:

- *handheld*: input types that make use of handheld controllers, such as smartphones and the wired trackpads used by Sony's SmartEyeglass and Epson's Moverio glasses.
- *touch*: non-handheld touch input, such as gesturing and tapping on body surfaces, and touch-sensing wearable devices (e.g. smart rings, watches, and glasses). These provide tactile feedback.
- *non-touch*: non-handheld, non-touch input, such as in-air gestures, head/body move-

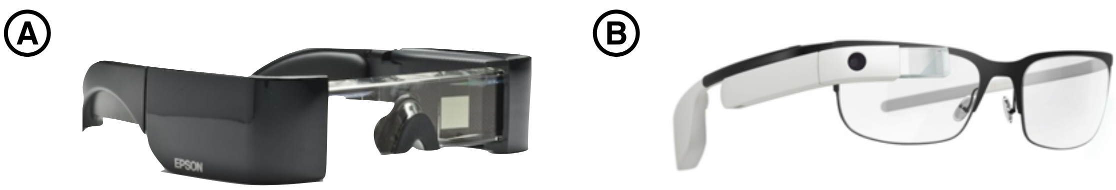

Figure 1.2: (A) Epson Moverio, (B) Google Glass.

ment, and voice recognition. These do not have tactile feedback.

We recruited 24 participants and asked them to wear the two form factors of smart glasses in a coffee shop. On-screen instructions prompted participants to perform each of the 17 common game control tasks using the 3 classes of input types. For each game control task, form factor, and input type, participants first explored all possible interactions they could think of, then reported the one they most preferred. After completing the 3 types of interactions for that task and form factor, they then rated their preferences for 3 interactions. Overall, each participant reported 102 interactions, for a total of 2448.

We collected quantitative and qualitative data through video analysis, preference ratings, and interviews. Our key observations are as follows:

- Participants significantly preferred non-handheld, non-touch interactions over handheld interactions (3.81 vs 3.68 on a 5-point Likert scale, p<0.01).

- For touch input without using handheld devices, users preferred interacting with their body surface over wearable devices (80% vs 20%), and the most frequently used body surface was the palm (51%).

- Participants preferred interactions that are more subtle due to concerns with social acceptance. Also, participants preferred using in-air gestures in front of the torso than in front of the face (63% vs 37%), even though those gestures were reported to be less intuitive and less precise.

- There is a significant mismatch between participants' preferred input methods and those supported by the current smart glasses. For example, less than 2% of the participants used voice and less than 2% of the participants used touch input on the smart glasses – which are Google Glass' two primary input methods. In addition, current cameras can only detect in-air gestures in front of users' faces, missing most 63% of the gestures performed.

The contribution of this paper are as follows: (1) the first quantitative and qualitative characterization of user-defined input for games on smart glasses, including a taxonomy,

(2) set of user-defined input for common game tasks, which is reflective of user behavior.
(3) insight into users' mental models when playing smart glasses games in a public space, and an understanding of implications for mobile input technology and user interface design. Our results will help designers create better smart glasses experience informed by user behavior.

Chapter 2

Related Work

2.1 Game Input

There are many previous works exploring new kinds of game controls and pushing the limit of game design. Vickers et al.[25] showed the possibility to use eye gestures as game inputs; Christian et al.[5] provided novel techniques for users to interact with games by head-gesture; Harada et al.[9] and Sporka et al.[24] both indicated that the voice input greatly expanded the scope of games that could be played hands-free and just counted on voice input; Baba et al.[3] presented a game prototype which treated skin contact as controller input; Nacke et al.[19] even considered using biofeedback (including EMG, EDA, EKG, RESP, TEMP) as game input methods; Hsu et al.[12] compared different game inputs, including head gestures, voice control, handheld controller, joysticks, eye winking and glass touchpad, for First-Person Shooter(FPS) games on smart glasses.

2.2 Mobile Input Technology

Some works related to mobile systems had defined designer-made input methods. These systems could be divided into two main categories, *touch* and *non-touch* inputs.

Harrison et al.[10] created *OmniTouch*, a wearable depth-sensing and projected system that enables interactive multitouch applications on any surface of the user's body. Moreover, *Skinput*[11], a technology that appropriates the human body for acoustic transmission, and allows the skin to be used as an input surface. Chan et al. presented FingerPad[4], a nail-mounted device that turns the tip of the index finger into a touchpad, allowing private and subtle interaction while on the move. Baudisch et al.[8] illustrated a concept of imaginary interface with sensing several gestures on the user's palms. Recently, Serrano et al.[23] explored the use of *Hand-to-Face* input to interact with head-worn displays(HWD) and provided a set of guidelines for developing effective Hand-to-Face interactions based on two main factors they found, social acceptability and cultural effect.

Kim et al.[15] developed a wrist-worn architecture, which supports discrete gesture recognition with reconstructing a 3D hand model in the air. Similarly, Jing et al.[13] implemented *Magic Ring*, a finger ring shaped input device using inertial sensors to detect subtle finger gestures; Colaço et al.[6] built a head-mounted display, *Mime*, sensing 3D gestures in front of the user's eyes.

2.3 Gestures in HCI

Gesture-based interfaces are already common in a variety of application domains such as gaming, virtual or augmented reality and mobile devices[14]. Aigner et al.[1] conducted a study of human preferences in usage of gesture types for HCI and indicated that, depending on the meaning of the gesture, there is preference in the usage of gesture types; Nielsen et al.[20] pointed out some important issues in choosing the set of gestures for the interface from a user-centred view such as the learning rate, ergonomics, and intuition; Grijincu et al.[7] presented a video-based gesture dataset and a methodology for annotating video-based gesture datasets; Recently, Piumsomboon et al.[21] have developed a user-defined gesture set for augmented reality applications. In our work, we focus on exploring relevant input gestures for gaming on smart glasses in public space.

2.4 User Elicitation Studies

User-elicitation studies are a specific type of participatory design methodology that involves end-users in the design of control-sets[18]. These studies had been used to design user interfaces of various types including multi-touch gestures on small and large surfaces[2, 27] and multi-modal interactions [18, 16]. There is also some evidence that user-defined control sets are more complete than those sets defined solely by experts[22, 27].

In a user-elicitation study, users were shown referents (an action's effects) and were asked to demonstrate the interactions that resulted in a given referent[27]. In this work, we draw upon the user-elicitation methodology to identify user expectations and suggestions for smart glass gaming.

Bibliography

[1] R. Aigner, D. Wigdor, H. Benko, M. Haller, D. Lindbauer, A. Ion, S. Zhao, and J. Koh. Understanding mid-air hand gestures: A study of human preferences in usage of gesture types for hci. *Microsoft Research TechReport MSR-TR-2012-111*, 2012.

[2] L. Anthony, Q. Brown, J. Nias, B. Tate, and S. Mohan. Interaction and recognition challenges in interpreting children's touch and gesture input on mobile devices. ITS'12, pages 225–234. ACM.

[3] T. Baba, T. Ushiama, R. Tsuruno, and K. Tomimatsu. Video game that uses skin contact as controller input. SIGGRAPH'07. ACM.

[4] L. Chan, R.-H. Liang, M.-C. Tsai, K.-Y. Cheng, C.-H. Su, M. Y. Chen, W.-H. Cheng, and B.-Y. Chen. Fingerpad: Private and subtle interaction using fingertips. UIST'13. ACM.

[5] S. Christian, J. Alves, A. Ferreira, D. Jesus, R. Freitas, and N. Vieira. Volcano salvation: Interaction through gesture and head tracking. CHI EA'14, pages 297–300. ACM.

[6] A. Colaço, A. Kirmani, H. S. Yang, N.-W. Gong, C. Schmandt, and V. K. Goyal. Mime: Compact, low power 3d gesture sensing for interaction with head mounted displays. UIST'13, pages 227–236. ACM.

[7] D. Grijincu, M. A. Nacenta, and P. O. Kristensson. User-defined interface gestures: Dataset and analysis. ITS'14, pages 25–34. ACM.

[8] S. Gustafson, C. Holz, and P. Baudisch. Imaginary phone: Learning imaginary interfaces by transferring spatial memory from a familiar device. UIST'11, pages 283–292. ACM.

[9] S. Harada, J. O. Wobbrock, and J. A. Landay. Voice games: Investigation into the use of non-speech voice input for making computer games more accessible. INTERACT'11, pages 11–29. Springer-Verlag.

[10] C. Harrison, H. Benko, and A. D. Wilson. Omnitouch: Wearable multitouch interaction everywhere. UIST'11, pages 441–450. ACM.

[11] C. Harrison, D. Tan, and D. Morris. Skinput: Appropriating the body as an input surface. CHI'10, pages 453–462. ACM.

[12] C.-Y. Hsu, Y.-C. Tung, H.-Y. Wang, S. Chyou, J.-W. Lin, and M. Y. Chen. Glass shooter: Exploring first-person shooter game control with google glass. ICMI'14, pages 70–71. ACM.

[13] L. Jing, Z. Cheng, Y. Zhou, J. Wang, and T. Huang. Magic ring: A self-contained gesture input device on finger. MUM'13, pages 39:1–39:4. ACM.

[14] M. Karam et al. A taxonomy of gestures in human computer interactions. 2005.

[15] D. Kim, O. Hilliges, S. Izadi, A. D. Butler, J. Chen, I. Oikonomidis, and P. Olivier. Digits: Freehand 3d interactions anywhere using a wrist-worn gloveless sensor. UIST'12, pages 167–176. ACM.

[16] H.-N. Liang, C. Williams, M. Semegen, W. Stuerzlinger, and P. Irani. User-defined surface+motion gestures for 3d manipulation of objects at a distance through a mobile device. APCHI'12, pages 299–308. ACM.

[17] Google Glass Mini Games. https://developers.google.com/glass/samples/mini-games.

[18] M. R. Morris. Web on the wall: Insights from a multimodal interaction elicitation study. ITS'12, pages 95–104. ACM.

[19] L. E. Nacke, M. Kalyn, C. Lough, and R. L. Mandryk. Biofeedback game design: Using direct and indirect physiological control to enhance game interaction. CHI'11, pages 103–112. ACM.

[20] M. Nielsen, M. Störring, T. B. Moeslund, and E. Granum. A procedure for developing intuitive and ergonomic gesture interfaces for hci. In *Gesture-Based Communication in Human-Computer Interaction*, pages 409–420. Springer, 2004.

[21] T. Piumsomboon, A. Clark, M. Billinghurst, and A. Cockburn. User-defined gestures for augmented reality. CHI'13, pages 955–960, New York, NY, USA. ACM.

[22] D. Pyryeskin, M. Hancock, and J. Hoey. Comparing elicited gestures to designer-created gestures for selection above a multitouch surface. ITS'12, pages 1–10. ACM.

[23] M. Serrano, B. M. Ens, and P. P. Irani. Exploring the use of hand-to-face input for interacting with head-worn displays. CHI'14, pages 3181–3190. ACM.

[24] A. J. Sporka, S. H. Kurniawan, M. Mahmud, and P. Slavík. Non-speech input and speech recognition for real-time control of computer games. Assets'06, pages 213–220. ACM.

[25] S. Vickers, H. Istance, and A. Hyrskykari. Performing locomotion tasks in immersive computer games with an adapted eye-tracking interface. *ACM Trans. Access. Comput.*, 5(1):2:1–2:33, Sept. 2013.

[26] J. O. Wobbrock, H. H. Aung, B. Rothrock, and B. A. Myers. Maximizing the guessability of symbolic input. CHI EA'05, pages 1869–1872. ACM.

[27] J. O. Wobbrock, M. R. Morris, and A. D. Wilson. User-defined gestures for surface computing. CHI'09, pages 1083–1092. ACM.