GlassMessaging: Supporting Messaging Needs During Daily Activities Using OST-HMDs

Nuwan Janaka
nuwanj@comp.nus.edu.sg
National University of Singapore
Singapore

Jie Gao
gaojie056@gmail.com
Singapore University of Technology and Design
Singapore

Lin Zhu
zhu-l20@mails.tsinghua.edu.cn
Tsinghua University
Beijing, China

Shengdong Zhao
Lan Lyu
Peisen Xu
zhaosd@comp.nus.edu.sg
lanlyu@andrew.cmu.edu
peisen.xu@cnrsatcreate.sg
National University of Singapore
Singapore

Maximilian Nabokow
mnabokow@cs.unc.edu
University of North Carolina at Chapel Hill
North Carolina, United States

Silang Wang
Yanch Ong
wangsilangCHANGE@gmail.com
ongyc1999@gmail.com
National University of Singapore
Singapore

ABSTRACT

The act of communicating with others during routine daily tasks is both common and intuitive for individuals. However, the hands- and eyes-engaged nature of present digital messaging applications makes it difficult to message someone amidst such activities. We introduce *GlassMessaging*, a messaging application designed for Optical See-Through Head-Mounted Displays (OST-HMDs). It facilitates messaging through both voice and manual inputs, catering to situations where hands and eyes are preoccupied. *GlassMessaging* was iteratively developed through a formative study identifying current messaging behaviors and challenges in common multitasking with messaging scenarios.

KEYWORDS

texting, messaging, heads-up computing, smart glasses, OST-HMD

ACM Reference Format:
Nuwan Janaka, Jie Gao, Lin Zhu, Shengdong Zhao, Lan Lyu, Peisen Xu, Maximilian Nabokow, Silang Wang, and Yanch Ong. 2023. *GlassMessaging*: Supporting Messaging Needs During Daily Activities Using OST-HMDs. In *The 2023 ACM Symposium on Spatial User Interaction (SUI '23), October 13–15, 2023, Sydney, NSW, Australia.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3607822.3618016

1 INTRODUCTION AND RELATED WORK

The proliferation of mobile devices has transformed our means of communication, making applications (henceforth referred to as *apps*) like WhatsApp, Telegram, Messenger, and WeChat commonplace [4]. However, using these apps during daily tasks, such as cooking or walking, is hindered by their design, which demands extensive visual and manual interaction. Research reveals that individuals often use messaging apps while multitasking, with 13% of messages sent on the move [1].

Given this context, we ask, *"How can we refine mobile messaging for effective communication during routine multitasking?"* This brings us to Optical See-Through Head-Mounted Displays (OST-HMDs or Augmented Reality Smart Glasses) [8], designed for hands-free usage and maintaining situational awareness [10, 14]. There remains a void in crafting interfaces tailored for OST-HMDs suited to daily multitasking, with current messaging apps for OST-HMDs primarily being derivatives of mobile phone apps (A notable exception is Google Glass XE (2013-2017) [5, 6] which is discussed in Appendix B).

The inclination to communicate while multitasking is evident in messaging app usage [4], fostering closeness and support [2]. Mobile phones, while supporting multitasking, can be hazardous in situations needing acute awareness, such as walking [7]. OST-HMDs appear promising due to their hands-free nature and enhanced situational awareness. Voice input stands out as a feasible hands-free technique for OST-HMDs, as other methods like head and gaze inputs might be less accurate or result in ergonomic strain [3, 11].

Consequently, we present *GlassMessaging* [9], a messaging application for OST-HMDs, iteratively designed post examining the prevailing needs, habits, challenges, and constraints users encounter while messaging and multitasking.

2 SYSTEM

To cater to hands-busy scenarios, we introduced **voice dictation** for text entry and **voice commands** for hands-free UI navigation. We also implemented a **ring mouse interaction** to allow for faster and more precise scrolling and selection while retaining **mid-air gestures** due to their "intuitive" touch-like content manipulation paradigm [9], as shown in Figure 1 (For apparatus and supported interactions, please see Appendix A).

SUI '23, October 13–15, 2023, Sydney, NSW, Australia

ACM ISBN 979-8-4007-0281-5/23/10.
https://doi.org/10.1145/3607822.3618016

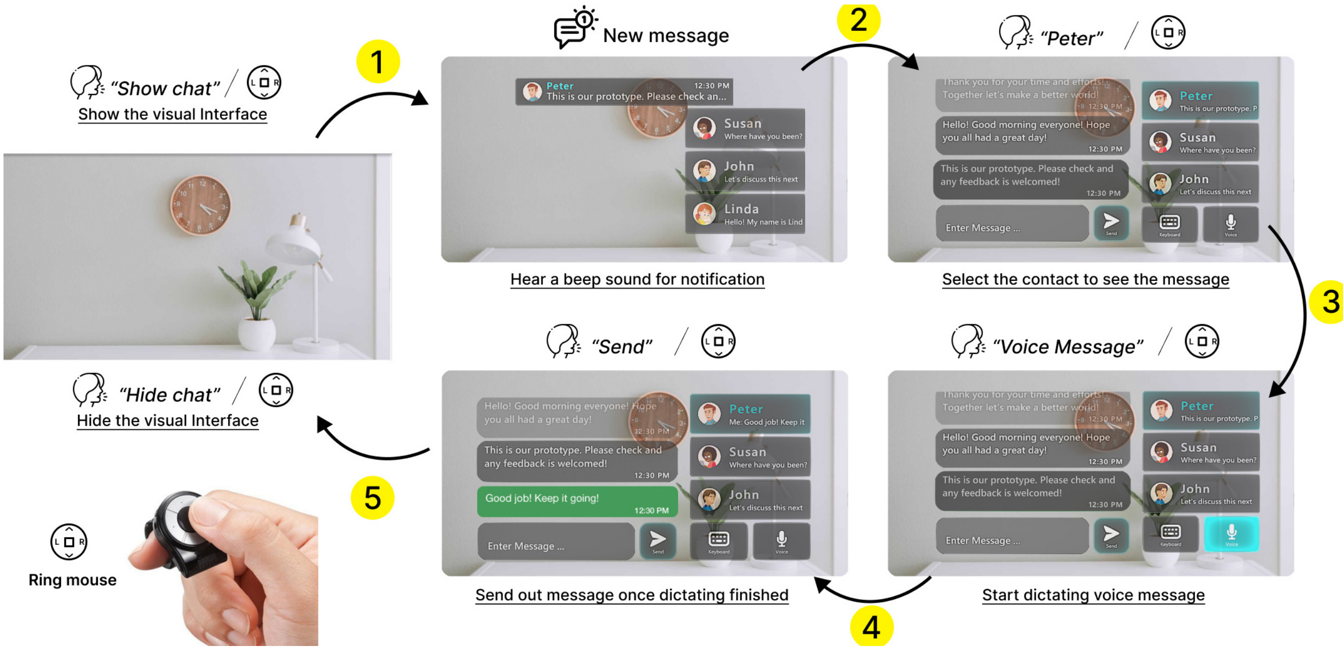

Figure 1: Steps for sending a message after receiving a notification. The user wears the OST-HMD and a ring mouse and sees the environment. The user (1) says 'SHOW CHAT', the interface is displayed, and a notification including the name of the sender and the sending time appears at the top of the view with a beep sound; (2) says the name of the contact (e.g., 'PETER'), and the system automatically navigates to the chat interface of the respective contact; (3) says 'VOICE MESSAGE' and dictates the message via voice. The system transcribes the user's utterances to text in real-time, displayed in the text entry box. Once the user stops speaking for a measured amount of time (silence gap), dictation turns off automatically; (4) says 'SEND', and the system sends the message; (5) says 'HIDE CHAT', and the full interface is hidden, restoring the full vision of the environment.

3 EVALUATION

To assess the effectiveness of *GlassMessaging*, we compared it to the Telegram application on mobile phones in a controlled study setting (N=16) in daily multitasking situations. Our findings [9] indicate that, even with the present technological constraints of the OST-HMD platform, *GlassMessaging* provided enhanced voice input access and enabled smoother interactions than phones. This resulted in a 33.1% reduction in response time and a 40.3% increase in texting speed. These findings underscore the significant potential of OST-HMDs as a meaningful complement to mobile phone-based messaging in multitasking scenarios.

However, there are several challenges to overcome before fully harnessing this platform's potential. For example, the use of *GlassMessaging* resulted in a 2.5% drop in texting accuracy, especially with complex texts. Moreover, current OST-HMDs have some inherent downsides (e.g., rudimentary hardware capabilities, unfamiliarity, limited interactions [8, 11]) when contrasted with the mature and extensively tested mobile phones currently available.

4 CONCLUSION AND FUTURE WORK

While multitasking with messaging is a frequent real-life activity, current mobile applications and platforms fall short in providing adequate support. We pinpointed two primary situational impediments (i.e., hands-busy and eyes-busy) arising from existing mobile platforms, which drove us to iteratively develop *GlassMessaging*, a messaging application tailored for OST-HMDs to address these shortcomings. We envision messaging on OST-HMDs as the forthcoming communication frontier, acting as a valuable adjunct to mobile phones during multitasking and driven forward by technological progress. To realize this vision, it is essential to re-conceptualize communication interfaces that align with OST-HMD affordances and to devise strategies to overcome potential situational challenges (e.g., privacy and social concerns with voice).

ACKNOWLEDGMENTS

We thank the volunteers who participated in our studies. This research is supported by the National Research Foundation, Singapore, under its AI Singapore Programme (AISG Award No: AISG2-RP-2020-016). It is also supported in part by the Ministry of Education, Singapore, under its MOE Academic Research Fund Tier 2 programme (MOE-T2EP20221-0010), and by a research grant #22-5913-A0001 from the Ministry of Education of Singapore. Any opinions, findings and conclusions, or recommendations expressed in this material are those of the author(s) and do not reflect the views of the National Research Foundation or the Ministry of Education, Singapore.

REFERENCES

[1] Agathe Battestini, Vidya Setlur, and Timothy Sohn. 2010. A large scale study of text-messaging use. In *Proceedings of the 12th international conference on Human computer interaction with mobile devices and services (MobileHCI '10)*.

Association for Computing Machinery, New York, NY, USA, 229–238. https://doi.org/10.1145/1851600.1851638

[2] Hyunsung Cho, Jinyoung Oh, Juho Kim, and Sung-Ju Lee. 2020. I Share, You Care: Private Status Sharing and Sender-Controlled Notifications in Mobile Instant Messaging. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW1 (May 2020), 1–25. https://doi.org/10.1145/3392839

[3] P R Cohen and S L Oviatt. 1995. The role of voice input for human-machine communication. *Proceedings of the National Academy of Sciences* 92, 22 (Oct. 1995), 9921–9927. https://doi.org/10.1073/pnas.92.22.9921

[4] David Curry. 2022. Most Popular Apps (2022). https://www.businessofapps.com/data/most-popular-apps/

[5] Google Glass. 2013. Google Glass - YouTube. https://www.youtube.com/user/googleglass

[6] Google Glass. 2013. Google Glass: How to use voice actions. https://www.youtube.com/watch?v=rv3KU0Yo5ZM

[7] Rami Hashish, Megan E. Toney-Bolger, Sarah S. Sharpe, Benjamin D. Lester, and Adam Mulliken. 2017. Texting during stair negotiation and implications for fall risk. *Gait & Posture* 58 (Oct. 2017), 409–414. https://doi.org/10.1016/j.gaitpost.2017.09.004

[8] Yuta Itoh, Tobias Langlotz, Jonathan Sutton, and Alexander Plopski. 2021. Towards Indistinguishable Augmented Reality: A Survey on Optical See-through Head-mounted Displays. *Comput. Surveys* 54, 6 (July 2021), 120:1–120:36. https://doi.org/10.1145/3453157

[9] Nuwan Janaka, Jie Gao, Lin Zhu, Shengdong Zhao, Lan Lyu, Peisen Xu, Maximilian Nabokow, Silang Wang, and Yanch Ong. 2023. GlassMessaging: Towards Ubiquitous Messaging Using OHMDs. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 7, 3 (Sept. 2023). https://doi.org/10.1145/3610931

[10] Nuwan Janaka, Xinke Wu, Shan Zhang, Shengdong Zhao, and Petr Slovak. 2022. Visual Behaviors and Mobile Information Acquisition. https://doi.org/10.48550/arXiv.2202.02748

[11] Lik-Hang Lee and Pan Hui. 2018. Interaction Methods for Smart Glasses: A Survey. *IEEE Access* 6 (2018), 28712–28732. https://doi.org/10.1109/ACCESS.2018.2831081

[12] Niora. 2023. Google Glass 3.0 - Review. https://www.niora.net/en/p/google_glass_3

[13] Wikipedia. 2023. Google Glass. https://en.wikipedia.org/w/index.php?title=Google_Glass&oldid=1139958030

[14] Shengdong Zhao, Felicia Tan, and Katherine Fennedy. 2023. Heads-Up Computing Moving Beyond the Device-Centered Paradigm. *Commun. ACM* 66, 9 (Aug. 2023), 56–63. https://doi.org/10.1145/3571722

A APPARATUS AND INTERACTIONS

We selected Microsoft HoloLens 2 (HL2), an OST-HMD with hand-tracking, voice commands, and world-scale positioning (2k resolution, 52° diagonal FoV), to develop *GlassMessaging*, our messaging app designed for OST-HMDs. A wireless ring mouse (Sanwa Supply 400-MA077) facilitated easy directional UI element selection (Figure 1). We developed *GlassMessaging* using Unity 3D, Mixed Reality Toolkit (MRTK 2.8), leveraging MRTK's built-in functions for mid-air gestures, voice inputs, virtual keyboard, and content stabilization. To simulate a realistic messaging experience, we implemented a virtual chat server using Python, running on a tablet computer connected to the HL2 via Wi-Fi, enabling bi-directional communication through a socket connection between the client and the server (see https://github.com/NUS-HCILab/GlassMessaging).

The output interface of *GlassMessaging* (as depicted in Figure 1) includes four main UI panels: notifications, contacts, chat messages, and voice/keyboard input panels, as well as audio outputs. These elements enable users to receive notifications, select contacts, and compose/send messages using both voice and keyboard inputs. The input interactions supported by this system are detailed in Table 1 and encompass voice commands, voice dictations, and manual inputs through ring interactions and mid-air hand gestures.

B *GLASSMESSAGING* VS. GOOGLE GLASS XE

Google Glass XE (2013-2017) [5, 6, 13], a discontinued product, supported heads-up messaging. Here, we distinguish between our application and Google Glass XE, showcasing our contributions from both practical and academic perspectives.

B.1 Google Glass XE (*GG*) interface

GG incorporated a default set of voice action commands for messaging [6]. Its lightweight and seamless design combined voice, head gestures, and touch gestures for inputs and an OST-HMD for output. To activate voice commands or send messages, users would utter "OK Glass" and "Send a message to", followed by the contact's name and message content. Users would respond to a message by saying "Reply" followed by their message content. Hence, *GG* provided an efficient method for sending and replying to *individual* messages.

B.2 Comparison

Table 2 depicts that both *GlassMessaging* (*GM*) and *GG* utilize voice input for text entry and navigation. Our study [9] validates voice input as an efficient tool, aligning with *GG*'s design. However, speech recognition affects the accuracy of *GM*, a challenge possibly shared by *GG* users. *GG*, while catering to immediate messaging requirements, had difficulty managing intricate conversations. In contrast, *GM* upholds modern standards, emphasizing context via features like full chat history and unread indicators. The display location differs too: *GG* showcased content above the line-of-sight, demanding attention shifts, while *GM*, leveraging advancements, positions content within the line-of-sight, employing opacity adjustments for awareness.

Interaction-wise, *GG* relied on head-gestures, while *GM* introduced a gamut of methods like ring and mid-air gestures, providing flexibility for multitasking. Ultimately, while both platforms serve heads-up messaging, their design nuances cater to different generational hardware and user demands. *GG*, tailored for earlier-generation glasses, prioritized real-time singular messages. *GM*, on the other hand, leverages advanced OST-HMDs, managing both immediate and layered messaging. A fusion of their strengths, such as integrating *GG*'s head-gesture system into *GM*, could potentially amplify user experience, especially in intricate messaging scenarios.

Table 1: Supported input interactions of *GlassMessaging*

Function	Mid-air gesture	Ring interaction	Voice command
Reveal interface		Click 'center' button for 1 second	Show chat
Hide interface		Click 'center' button for 1 second	Hide chat
Open the chat related to notification	Press on notification		Open notification
Open the chat with the contact, <NAME>	Press on contact <NAME>	Click 'up'/'down' button to navigate	<NAME>
Activate voice dictation	Press on 'voice' button	Click 'right' button to navigate + click 'center' button to activate	Voice message
Send the message	Press on 'send' button		Send
Open the virtual keyboard	Press on 'keyboard' button		Open keyboard
Close the virtual keyboard	Press anywhere on the interface	Click any button	Close keyboard
Go to the topmost contact	Scroll up using the finger	Click and hold the 'up' button	Scroll to the top
Go to closest top contact	Scroll up using the finger	Click 'up' button	Scroll up
Go to closest bottom contact	Scroll down using the finger	Click 'down' button	Scroll down
Start voice dictating to message received contact			Reply (Open notification + Voice message)
Start voice dictation to the contact, <NAME>			Text <NAME> (<NAME> + Voice message)

Table 2: A comparison between *GlassMessaging* and Google Glass XE. LoS stands for Line of Sight, and FoV represents Field of View. Note: This list is not exhaustive and based on public online resources [5, 12, 13], as Google Glass XE has been discontinued since 2017.

Features	*GlassMessaging* with HL2	Google Glass XE (*GG*)
Interactions	Voice, Ring-mouse, Mid-air	Voice, Touchpad (on the right temple), Head gestures
Text entry	Voice, Mid-air keyboard	Voice
Display	Binocular, Higher-Resolution (2048x1080 px per eye) Larger-FoV (30° horizontal)	Monocular, Lower-resolution (640x360 px, right eye) Smaller-FoV (13° horizontal)
Chat history	Shows last 3 messages and 3 contacts	Shows last message and last contact
Chat position Notification position Contact position	LoS (middle-center) Above LoS (top-center) Right of LoS (middle-right)	Above LoS (top-center), Manual switching between each UI
UI opacity	Increased for new messages	Fixed
UI access	On-demand (using voice or ring-mouse)	On-demand (by looking up or using voice)