# GesMessages: Using Mid-air Gestures to Manage Notifications

Xiang Li
xl529@cam.ac.uk
University of Cambridge
Cambridge, United Kingdom

Yuzheng Chen
yuzhengchen@hkust-gz.edu.cn
Hong Kong University of Science and Technology (Guangzhou)
Guangzhou, China

Xiaohang Tang
xiaohangtang@vt.edu
Virginia Tech
Blacksburg, VA, USA

## ABSTRACT

This paper introduces GesMessages, an innovative mid-air interactive application that uses simple gestures to manage real-time message notifications on laptops and large displays. Leveraging cameras on computers or smart devices, the application offers three distinct gestures: expanding notifications for immediate attention, hiding non-urgent messages, and deleting spam messages. We present the technical setup and system design. Additionally, we explore potential applications in context-awareness systems, contributing to gestural interaction research. Our work fosters a deeper understanding of mid-air interaction's impact on message management and future interactive systems.

## CCS CONCEPTS

• **Human-centered computing** → **Ubiquitous and mobile computing systems and tools**; **Gestural input**.

## KEYWORDS

Augmented reality, gesture input, message notification

**ACM Reference Format:**
Xiang Li, Yuzheng Chen, and Xiaohang Tang. 2023. GesMessages: Using Mid-air Gestures to Manage Notifications. In *The 2023 ACM Symposium on Spatial User Interaction (SUI '23), October 13–15, 2023, Sydney, NSW, Australia.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3607822.3618023

## 1 INTRODUCTION

Mid-air interaction, characterized by kinaesthetic engagement [10], has revolutionized human-computer interaction (HCI) by providing an intuitive and immersive experience through the use of human gestures and movements [12, 16, 17]. Although successfully implemented in Augmented Reality (AR) and Virtual Reality (VR) headsets, its adoption in conventional computing devices, like laptops, remains limited. However, the constant accessibility of information across diverse contexts [6, 7, 9, 13] presents a challenge. Interruptions from email and message notifications can disrupt focus during activities, even during moments of leisure, such as watching a movie on a laptop. Despite advancements in mid-air interaction techniques, users still resort to traditional input devices, like the mouse, to manage notifications alongside their primary tasks.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).
*SUI '23, October 13–15, 2023, Sydney, NSW, Australia*
© 2023 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-0281-5/23/10.
https://doi.org/10.1145/3607822.3618023

In this paper, we present a mid-air interactive application tailored for real-time message notification management on laptops and large displays to address this challenge. Leveraging intuitive gestures [14], our application empowers users to seamlessly handle pop-up notifications without sacrificing their focus. Utilizing computer or smart device cameras, our system offers three key gestures: (1) expanding notifications for immediate attention, (2) hiding non-urgent messages for later handling, and (3) deleting spam messages for an uncluttered interface. These gestures are designed to be ergonomic and effortless for users.

This paper presents the technical setup and system design of our. Furthermore, we explore potential enhancements to its usability and discuss its integration with context-awareness systems, particularly in the context of everyday information acquisition. Our work contributes to the ongoing exploration of effective HCI techniques and aims to improve the digital experience for users.

## 2 RELATED WORK

Designing effective mid-air gesture sets is a challenging task [5, 8, 20, 21]. Prior research emphasizes the importance of creating gestures that are easy to act reliably [5], consider reliability and learnability [2], and are immediately usable [21]. Numerous studies explore mid-air interaction across different scenarios and platforms. Some focus on VR/AR environments, investigating mid-air interaction for web browsing [4], mid-air typing schemes [22], and target selection techniques for mobile VR [11]. Others explore 3D manipulation operations using mid-air interaction [19]. Beyond VR/AR setups, mid-air interaction is studied in diverse contexts. It includes browsing hierarchical information on interactive public displays [1], pan-and-zoom techniques for wall-sized displays [18], and a mid-air bi-manual interaction approach for gestural text input on laptops [3]. These explorations contribute to the advancement of mid-air interaction technologies in various real-world scenarios.

## 3 SYSTEM

### 3.1 Technical Setup

We used Unity 2021.2.13f1 as the platform for detecting hand gestures. The recognition algorithm and the parameters collection were based on the MediaPipe, which is a framework for building pipelines to perform inference over arbitrary sensory data [15].

### 3.2 Design

We only provide three simple features for expanding the pop-up notification, hiding the messages, and deleting the messages, as a work-in-progress prototype. When a new message pops up, the user can hold one hand on the top right for about a second and the details of the message will be expanded (see Figure 1 (1)). When

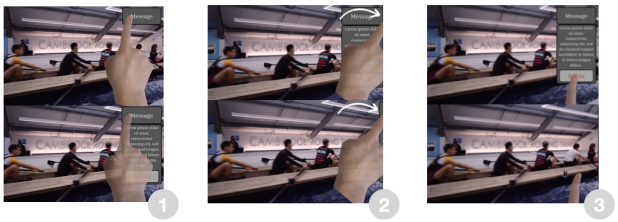

**Figure 1: Interaction design of GesMessages: (1) to expand the details of a new message, the user can hold one hand on the top right corner for about a second. (2) to dismiss a new message notification, the user can drag the message to the right edge of the screen. The message will be hidden once it is removed from the screen boundaries. (3) to delete a new message, the user can hover their index finger over it and then drag it to the trash bin at the bottom of the screen.**

a new pop-up notification comes, the user can drag the message and move it to the right of the screen. And once this message is moved out of the screen boundaries, the message notification will be hidden for the user (see Figure 1 (2)). If the new message is spam, the user can let his/her index finger hover right on the message, and then, they can drag the message to the trash bin (at the button of the screen) (see Figure 1 (3)).

## 4 FUTURE WORK AND CONCLUSION

In this paper, we present a mid-air interactive application, GesMessages, that can manage real-time message notifications with only a few simple gestures. It can be used for laptops and big displays. We provide three gestures based on the camera of a computer or other smart device to detect and adjust the pop-up messages, i.e., expand the pop-up notification, hide the message, and delete the message. We articulate the technical setup of this application, its design, and present preliminary findings, and mention our future work in the form of two themes: (a) the usability and efficiency of our application, and (b) how it can be used for the context-awareness systems, especially for the everyday information acquisition. Ultimately, we hope that our subsequent work can contribute to the understanding of gestural interaction in information acquisition and context-awareness systems.

## ACKNOWLEDGMENTS

Xiang Li is supported by the China Scholarship Council (CSC) International Cambridge Scholarship (No. 202208320092).

## REFERENCES

[1] Christopher Ackad, Andrew Clayphan, Martin Tomitsch, and Judy Kay. 2015. An in-the-wild study of learning mid-air gestures to browse hierarchical information at a large interactive public display. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. 1227–1238.

[2] Thomas Baudel and Michel Beaudouin-Lafon. 1993. Charade: remote control of objects using free-hand gestures. *Commun. ACM* 36, 7 (1993), 28–35.

[3] Garrett Benoit, G Michael Poor, and Alvin Jude. 2017. Bimanual Word Gesture Keyboards for Mid-air Gestures. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. 1500–1507.

[4] Assam Boudjelthia, Sofeem Nasim, Janne Eskola, Joshua Muyiwa Adeegbe, Oula Hourula, Simon Klakegg, and Denzil Ferreira. 2018. Enabling mid-air browser interaction with leap motion. In *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers*. 335–338.

[5] Debaleena Chattopadhyay and Davide Bolchini. 2015. Motor-intuitive interactions based on image schemas: Aligning touchless interaction primitives with human sensorimotor abilities. *Interacting with Computers* 27, 3 (2015), 327–343.

[6] Karen Church, Mauro Cherubini, and Nuria Oliver. 2014. A large-scale study of daily information needs captured in situ. *ACM Transactions on Computer-Human Interaction (TOCHI)* 21, 2 (2014), 1–46.

[7] Shakiba Davari, Feiyu Lu, and Doug A Bowman. 2022. Validating the Benefits of Glanceable and Context-Aware Augmented Reality for Everyday Information Access Tasks. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 436–444.

[8] Nigel Davies, Sarah Clinch, and Florian Alt. 2014. Pervasive displays: understanding the future of digital signage. *Synthesis Lectures on Mobile and Pervasive Computing* 8, 1 (2014), 1–128.

[9] David Dearman, Melanie Kellar, and Khai N Truong. 2008. An examination of daily information needs and sharing opportunities. In *Proceedings of the 2008 ACM conference on Computer supported cooperative work*. 679–688.

[10] Maiken Hillerup Fogtmann, Jonas Fritsch, and Karen Johanne Kortbek. 2008. Kinesthetic interaction: revealing the bodily potential in interaction design. In *Proceedings of the 20th Australasian conference on computer-human interaction: designing for habitus and habitat*. 89–96.

[11] Akira Ishii, Takuya Adachi, Keigo Shima, Shuta Nakamae, Buntarou Shizuki, and Shin Takahashi. 2017. FistPointer: target selection technique using mid-air interaction for mobile VR environment. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. 474–474.

[12] Panayiotis Koutsabasis and Panagiotis Vogiatzidakis. 2019. Empirical research in mid-air interaction: A systematic review. *International Journal of Human–Computer Interaction* 35, 18 (2019), 1747–1768.

[13] Xiang Li, Yuzheng Chen, Rakesh Patibanda, and Florian 'Floyd' Mueller. 2021. vrCAPTCHA: Exploring CAPTCHA Designs in Virtual Reality. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems (CHI EA '21)*. Association for Computing Machinery, New York, NY, USA, 1–4. https://doi.org/10.1145/3411763.3451985

[14] Xiang Li, Yuzheng Chen, and Xiaohang Tang. 2022. GesPlayer: Using Augmented Gestures to Empower Video Players. In *Companion Proceedings of the 2022 Conference on Interactive Surfaces and Spaces* (Wellington, New Zealand) *(ISS '22)*. Association for Computing Machinery, New York, NY, USA, 4–8. https://doi.org/10.1145/3532104.3571456

[15] Camillo Lugaresi, Jiuqiang Tang, Hadon Nash, Chris McClanahan, Esha Uboweja, Michael Hays, Fan Zhang, Chuo-Ling Chang, Ming Guang Yong, Juhyun Lee, et al. 2019. Mediapipe: A framework for building perception pipelines. *arXiv preprint arXiv:1906.08172* (2019).

[16] Florian 'Floyd' Mueller, Rakesh Patibanda, Richard Byrne, Zhuying Li, Yan Wang, Josh Andres, Xiang Li, Jonathan Marquez, Stefan Greuter, Jonathan Duckworth, et al. 2021. Limited control over the body as intriguing play design resource. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–16.

[17] Florian 'Floyd' Mueller, Nathan Semertzidis, Josh Andres, Joe Marshall, Steve Benford, Xiang Li, Louise Matjeka, and Yash Mehta. 2023. Towards Understanding the Design of Intertwined Human-Computer Integrations. *ACM Trans. Comput.-Hum. Interact.* (apr 2023). https://doi.org/10.1145/3590766 Just Accepted.

[18] Mathieu Nancel, Julie Wagner, Emmanuel Pietriga, Olivier Chapuis, and Wendy Mackay. 2011. Mid-air pan-and-zoom on wall-sized displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 177–186.

[19] Peng Song, Wooi Boon Goh, William Hutama, Chi-Wing Fu, and Xiaopei Liu. 2012. A handle bar metaphor for virtual object manipulation with mid-air interaction. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 1297–1306.

[20] Robert Walter, Gilles Bailly, and Jörg Müller. 2013. StrikeAPose: revealing mid-air gestures on public displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 841–850.

[21] Robert Walter, Gilles Bailly, Nina Valkanova, and Jörg Müller. 2014. Cuenesics: using mid-air gestures to select items on interactive public displays. In *Proceedings of the 16th international conference on Human-computer interaction with mobile devices & services*. 299–308.

[22] Hao Zhang, Yafeng Yin, Lei Xie, and Sanglu Lu. 2020. Airtyping: A mid-air typing scheme based on leap motion. In *Adjunct Proceedings of the 2020 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2020 ACM International Symposium on Wearable Computers*. 168–171.