SenseBot: Supporting Embodied Remote Communication through AR-enabled Social Robot

Hao Jin*
jin.ha@northeastern.edu
Northeastern University
Boston, Massachusetts, USA

Lyumanshan Ye*
yelyumanshan@sjtu.edu.cn
Shanghai Jiao Tong University
Min hang, Shanghai, China

Qiao Jin
jin00122@umn.edu
University of Minnesota
Minneapolis, MN, Minnesota, USA

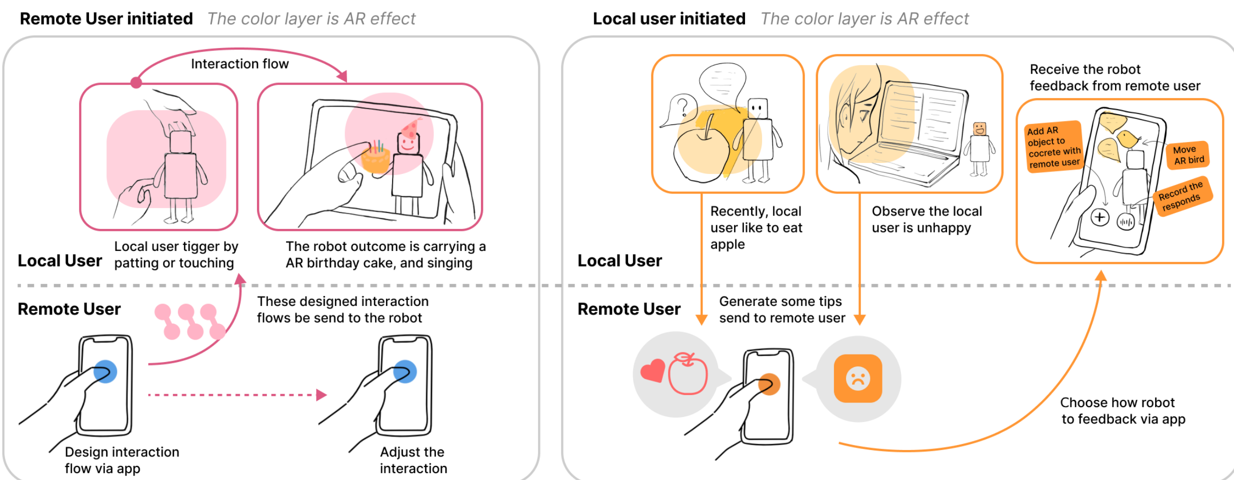

Figure 1: The overview charts shows how remote user interact with local user through AR-enabled social robot

Abstract

Embodied communication, which uses physical cues to convey meaning and emotions, is essential for building social connections. In this work, we introduce SenseBot, an augmented reality (AR)-enabled robotic representative designed to facilitate embodied communication. SenseBot acts as a local agent for remote users, allowing them to engage more interactively and meaningfully with people on-site.

CCS Concepts

• **Human-centered computing → Ubiquitous and mobile devices**; **Mixed / augmented reality**; **Web-based interaction**; **Collaborative interaction**.

Keywords

Robot, AR, social connection, remote communication, embodied communication

*Both authors contributed equally to this work.

UIST Adjunct '24, October 13–16, 2024, Pittsburgh, PA, USA

ACM ISBN 979-8-4007-0718-6/24/10
https://doi.org/10.1145/3672539.3686734

ACM Reference Format:
Hao Jin, Lyumanshan Ye, and Qiao Jin. 2024. SenseBot: Supporting Embodied Remote Communication through AR-enabled Social Robot. In *The 37th Annual ACM Symposium on User Interface Software and Technology (UIST Adjunct '24), October 13–16, 2024, Pittsburgh, PA, USA.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3672539.3686734

1 Introduction and Related Work

Social connections are crucial for health and well-being, reducing loneliness and promoting longevity. However, maintaining these bonds can be difficult without physical contact [3]. Embodied communication—using physical cues and gestures to convey meanings and emotions—becomes important in remote interactions. This approach is particularly effective for children, as it helps capture their attention and maintain their engagement [2], thus strengthens emotional bonds with their remote family members (e.g., siblings [5], grandparents and grandchildren [9, 10], children and parents [6]). However, most studies focus on symmetric, real-time communication, which may not always be feasible due to differing schedules, technological comfort levels, or device access. This highlights the need for more flexible communication systems to ensure meaningful connections across generations. For example, Haptic Bubble [1] enables playful interaction through a combination of physical and virtual components in an asymmetric setting, where children use physical devices while remote adults use a mobile application. This setup can be adapted for intergenerational communication, fostering emotional connections and engagement across different age groups [11].

Social robots are increasingly used in children's companionship by simulating human behavior and emotions [3, 7, 11]. Combining AR with robots enhances engagement through dynamic, real-time visual feedback. For example, AR Math [4] and user-generated AR [4, 8] shows that combining virtual systems with physical blocks provides engaging, intuitive learning with real-time feedback, enhancing interaction between children and social robots. To address these limitations, our system integrates AR technology into a social robot, providing user-defined visual cues and an immersive experience. By incorporating asymmetric design, we aim to strengthen emotional connections between children and their remote family members, blending flexibility with richer emotional expression.

2 System Design

We proposed SenseBot, an AR-based social robot designed to enhance remote inter-generational communication between children and distant family members (e.g., siblings, parents, grandparents). SenseBot serves as both a companion and an on-site representative for remote users, facilitating embodied interactions. With an asymmetric design—where remote and local users have different roles and interaction capabilities—SenseBot allows for flexible communication, accommodating varying schedules and contexts.

This section outlines the components for remote and local users in the SenseBot system and provides an overview of the embodied and AR interactions it supports.

2.1 Remote Client

The remote users mainly use a application to defined the AR and embodied interaction of SenseBot for local users. The application allows remote users to design their local agent, send embodied social interactions or be notified when received the message from local client. There are three different components involved in the SenseBot's application:

- **Customize robot's appearance:** First, remote users can design the robot's AR appearance as their communication agent. Robot (has head and body, attached with an AR image target) is customized with remote user's headshot and clothes. They can preview the robot's AR effects through the application. It doesn't need to be human-like, can be object, animal or cartoon characters.
- **Define trigger-outcome interaction flow:** Remote users can design how local users trigger the robot, specifying actions, trigger points on the robot, and the resulting embodied response or AR effect (see fig.1). These custom flows are sent to the robot via the app.
- **Receive notifications and send interactions:** Remote users receive lightweight notifications about the local user's status as observed by SenseBot (e.g., unhappy today, likes apples, see fig.1). They can then use the app to send asynchronous robot or AR interactions in response.

2.2 Local Client

Local users directly interact with robot, and they use an mobile application to see the AR effect of robot.

- **AR effects display:** Local users can touch, pat, wave hand(or other physics action)to trigger remote user-defined AR visual effects. They can use the application on their smartphone or pad, turn on the camera, and see the AR effect on the robot.
- **Audio feedback:** SenseBot can play audio, including music or natural speech, and engage in conversations using large language models.
- **Facial emotion detection:** The system uses AI-driven facial emotion recognition to detect the local user's emotional state. Based on detected emotions or surrounding objects, the robot generates soothing voice responses via LLMs and sends relevant notifications to the remote user's application.
- **Contextual information monitoring:** We plan to use Grove sensors, including the Tilt Switch, Temperature Sensor, Rotary Angle Sensor, and Light Sensor v1.2.
- **Physical interactions detection:** The system will utilize Grove sensors such as the Vibration Sensor (SW-420), Mini PIR Motion Sensor, Loudness Sensor, and Ultrasonic Ranger to detect physical interactions from the local user.

3 Workflow

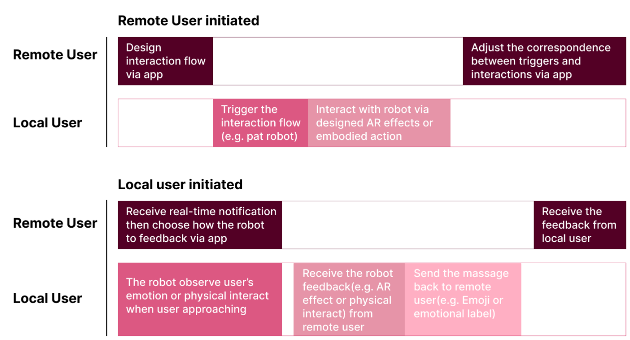

Figure 2: The interaction flow charts of remote user initiated and local user initiated communication.

In remote user-initiated interactions, the remote user designs interaction flows via the app, defining triggers like "pat my head" and corresponding AR effects or robot actions. The local user activates these flows through physical actions, prompting the robot to perform tasks like displaying an AR cake, an AR bird, or dancing. The local user can interact using the AR camera, and the robot can also provide audio responses, such as music or spoken messages.

In local user-initiated interactions, the robot uses AI vision to observe the user's emotions or actions. If the user appears unhappy, the robot notifies the remote user, who can then select a feedback method via the app, such as AR effects or physical interactions. The robot executes the chosen interaction and communicates the results to the local user. The system also allows the remote user to customize the robot's appearance and define new interaction flows.

These asynchronous workflows (see fig.2) ensure smooth communication between remote and local users, supporting both synchronous and asynchronous interactions to accommodate different schedules and enhance the user experience.

References

[1] Stuti Arora, Qiao Jin, and Svetlana Yarosh. 2023. Exploring Embodied Approaches for Large Age Gap Sibling Communication through Technology Probes. In *Proceedings of the 22nd Annual ACM Interaction Design and Children Conference* (Chicago, IL, USA) *(IDC '23)*. Association for Computing Machinery, New York, NY, USA, 635–640. https://doi.org/10.1145/3585088.3593892

[2] Anna Sophia Calmbach, Sophie Kunz, Alice C Haynes, and Jürgen Steimle. 2024. ChiParCo and ParChiCo: Connecting Children and Parents Remotely With Tailored Tangible Communication Tools. In *Proceedings of the 23rd Annual ACM Interaction Design and Children Conference* (Delft, Netherlands) *(IDC '24)*. Association for Computing Machinery, New York, NY, USA, 424–435. https://doi.org/10.1145/3628516.3655803

[3] Verena Fuchsberger, Janne Mascha Beuthel, Philippe Bentegeac, and Manfred Tscheligi. 2021. Grandparents and Grandchildren Meeting Online: The Role of Material Things in Remote Settings. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 478, 14 pages. https://doi.org/10.1145/3411764.3445191

[4] Qiao Jin, Danli Wang, Xiaozhou Deng, Nan Zheng, and Steve Chiu. 2018. AR-maze: a tangible programming tool for children based on AR technology. In *Proceedings of the 17th ACM Conference on Interaction Design and Children* (Trondheim, Norway) *(IDC '18)*. Association for Computing Machinery, New York, NY, USA, 611–616. https://doi.org/10.1145/3202185.3210784

[5] Qiao Jin, Ye Yuan, and Svetlana Yarosh. 2023. Socio-technical Opportunities in Long-Distance Communication Between Siblings with a Large Age Difference. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 94, 15 pages. https://doi.org/10.1145/3544548.3580720

[6] Bumsoo Kang, Seungwoo Kang, and Inseok Hwang. 2021. MomentMeld: AI-augmented Mobile Photographic Memento towards Mutually Stimulatory Inter-generational Interaction. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 479, 16 pages. https://doi.org/10.1145/3411764.3445688

[7] Zhao Liu, Ernuo Cheng, Xinyuan Zhang, and Xipei Ren. 2022. LUNOST: Connected Tangible Messengers for Enhancing Off-Site ParentTeenager Relationships. In *Proceedings of the Ninth International Symposium of Chinese CHI* (Online, Hong Kong) *(Chinese CHI '21)*. Association for Computing Machinery, New York, NY, USA, 127–132. https://doi.org/10.1145/3490355.3490372

[8] Ryo Suzuki, Adnan Karim, Tian Xia, Hooman Hedayati, and Nicolai Marquardt. 2022. Augmented Reality and Robotics: A Survey and Taxonomy for AR-enhanced Human-Robot Interaction and Robotic Interfaces. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 553, 33 pages. https://doi.org/10.1145/3491102.3517719

[9] Torben Wallbaum, Andrii Matviienko, Swamy Ananthanarayan, Thomas Olsson, Wilko Heuten, and Susanne C.J. Boll. 2018. Supporting Communication between Grandparents and Grandchildren through Tangible Storytelling Systems. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3173574.3174124

[10] Xiaoying Wei, Yizheng Gu, Emily Kuang, Xian Wang, Beiyan Cao, Xiaofu Jin, and Mingming Fan. 2023. Bridging the Generational Gap: Exploring How Virtual Reality Supports Remote Communication Between Grandparents and Grandchildren. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 444, 15 pages. https://doi.org/10.1145/3544548.3581405

[11] Ye Yuan, Peter Genatempo, Qiao Jin, and Svetlana Yarosh. 2024. Field Trial of a Tablet-based AR System for Intergenerational Connections through Remote Reading. *Proc. ACM Hum.-Comput. Interact.* 8, CSCW1, Article 205 (apr 2024), 28 pages. https://doi.org/10.1145/3653696