Demo: Towards Peripheral Awareness of Remote Family Member's Context Using Self-mobile Robotic Avatars

Bumsoo Kang[1], Inseok Hwang[2], Jinho Lee[2], Seungchul Lee[1]
Taegyeong Lee[1], Youngjae Chang[1], Min Kyung Lee[3]
[1]KAIST, [2]IBM, [3]Carnegie Mellon University
[1]{bumsoo, seungchul, tglee, yjchang}@nclab.kaist.ac.kr, [2]{ihwang, leejinho}@us.ibm.com, [3]mklee@cs.cmu.edu

ACM Reference Format:
Bumsoo Kang, Inseok Hwang, Jinho Lee, Seungchul Lee, Taegyeong Lee, Youngjae Chang, Min Kyung Lee. 2019. Demo: Towards Peripheral Awareness of Remote Family Member's Context Using Self-mobile Robotic Avatars. In *The 17th Annual International Conference on Mobile Systems, Applications, and Services (MobiSys '19), June 17–21, 2019, Seoul, Republic of Korea.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3307334.3328578

1 INTRODUCTION

Real-time remote interaction has become easier and richer powered by recent advances in mobile computing and communication. A number of research have been explored on enriching family interaction by augmenting an interaction channel with asynchronous communication [6] or additional sensory stimuli [5]. However, it is still far from achieving a sense of living together for family members involuntarily living apart, especially in context-aware impromptu interaction. For families living together, it is trivial to naturally perceive behavioral and situational contexts of the other and initiate a relevant interaction intuitively. For example, a wife starts a casual chat with asking her husband what he is going to cook when she sees him going to the kitchen or hears a simmering sound.

We envision that co-present robotic avatar mirroring a remote family member's activity in the one's local living space enables people to naturally aware the other's context without intention, namely peripheral awareness. We believe that such peripheral awareness could bring many pseudo living-together experiences, including context-aware impromptu interaction. As an initial attempt towards peripheral awareness, we focus on intelligent simultaneous positioning of the robotic avatar from a human point-of-view.

We demonstrate our initial prototype, HomeMeld[1], a self-mobile robotic system serving as a real-time, co-present avatar to provide peripheral awareness of the other's in-home context [7]. Figure 1 shows HomeMeld operating at two distant homes, the wife's home and the husband's home at the same time. The wife is sitting at her desk and talking to her husband's avatar locating at her couch and facing her. Conversely, the husband is sitting at his couch and talking to his wife's avatar locating at his desk and facing him.

[1]Our video is available at: https://goo.gl/avF5Ji

MobiSys '19, June 17–21, 2019, Seoul, Republic of Korea

ACM ISBN 978-1-4503-6661-8/19/06.
https://doi.org/10.1145/3307334.3328578

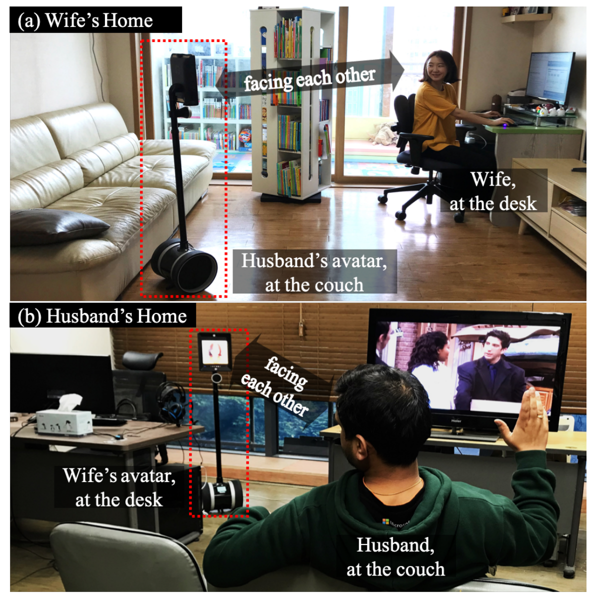

Figure 1: HomeMeld helps family members involuntarily living apart to naturally aware the other's context.

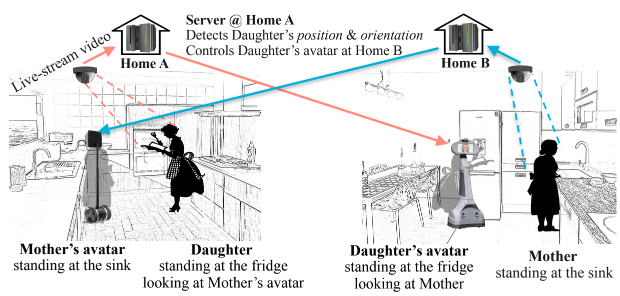

Figure 2: An operational overview of HomeMeld

HomeMeld is built on top of Double 2 [1], a commercial telepresence robot hardware, and YOLOv3 [8], a DNN-based object detection model. In this demo, we present the motivation behind our work, the end-to-end operation of HomeMeld, and the demonstration plan to convey our vision of pseudo living-together experiences across two mock-up living spaces.

2 HOMEMELD

HomeMeld is a self-mobile robotic avatar system with full end-to-end autonomous operating features, including: 1) localizing the person and the avatar, 2) mapping two heterogeneous floor plans, and 3) finding an optimal path to drive the avatar for the given origin, destination, and floor plan. Figure 2 shows an operational overview of the whole system consisting of a ceiling mounted camera, a telepresence robot, and a server module for each home.

Technical challenges in realizing HomeMeld include human detection, heterogeneous floor mapping, and synchronized robot navigation. The full paper [7] presents the motivating study of HomeMeld and more details on following technical challenges.

2.1 Human detection

The first step is to detect the location and orientation of the user. We use a ceiling-mounted 360° camera [2] in favor of obstruction-free human detection. Since detecting person from a top view has not been frequently addressed in existing object detection models, we created our custom dataset and trained the model. From the top view, a person's appearance is symmetric with respect to the center of the camera view. We benefit from this property and a green-screening technique to boost our handful volume of initial data to a practically large size to train a neural network model with. To be specific, we diversified our initial dataset by screening participants, randomly rotating screened images around the center of the camera view, and synthesizing them with background images.

2.2 Heterogeneous floor mapping

Reproducing a remote user's location and orientation in the local user's home is important towards realizing HomeMeld. A challenge to this is the spatial heterogeneity between the two homes, e.g., different floor plans and furniture arrangements. We define a notion of *functionally equivalent location and orientation* which is computed with respect to one or more prominent in-home objects (i.e., functional object). Being close to the object and looking at the object are strong indicators that the user is using and interacting with the object. In this light, we reproduce remote user's contexts with location and orientation with respect to the functional objects in local user's home. Our mapping approach (i.e., object-oriented mapping) partitions the given home space into several regions each of which encloses a functional object (e.g., Voronoi diagram [3]). Then a point in Home A is mapped into a point in Home B by finding an equivalent polar coordinate of the mapped point with respect to the functional object being considered.

2.3 Synchronized robot navigation

Synchronizing the arrival time at the destination is an important issue to timely convey user's activity update upon arrival. The difficulty comes from the fact that the robot is slower than a human, especially in rotations and accelerations. Therefore, we add a destination prediction feature on top of A* path optimization [4]. Based on the information of the user's real-time location and orientation, we attempt to predict his/her destination. We select the closest functional object in the user's moving direction, then find the optimal path from the current avatar's position to the predicted destination (i.e., the closest functional object). This approach helps the avatar arrive at the destination faster because it navigates along the optimized path to the destination rather than the tracing every transient functionally equivalent location of the moving person.

Figure 3: Example settings of the demo

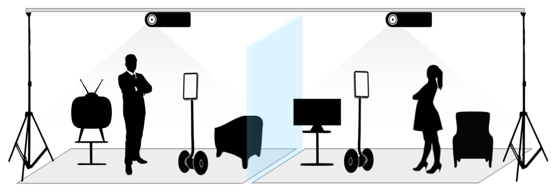

Figure 4: Planned demo setup with two separate spaces

3 DEMONSTRATION

We will separate the demo space into two bedroom-sized spaces and decorate each room to look like a living space by placing home furniture. Figure 3 shows example settings. Each space will include a 360° camera and a telepresence robot. Cameras will be placed on the home-studio like camera stand because it is not possible to put cameras on the 5m-high ceiling of the demo space (see Figure 4). To ensure smooth demo without potential network troubles, we will set up our server locally where the human detection model runs.

We plan to set up our demo that audience themselves can play and experience. Two people will be able to participate at a time; each of them moves around in one's own space with a telepresence robot that reflects the other participant across the wall. The demo space will be separated by a transparent wall to let participants be able to directly see their own avatars' movements in the other room, and intuitively understand our system's key features and operation. We will also prepare simulators for heterogeneous floor mapping to let people see how the mapping works while waiting for participation. We expect that participants will have experiences on end-to-end operation of HomeMeld, including human detection, heterogeneous floor mapping, and synchronized robot navigation.

4 ACKNOWLEDGEMENTS

This work was supported by the National Research Foundation of Korea (NRF) funded by the Ministry of Science and ICT (NRF-2017R1A2B3010504 and NRF-2017M3C4A7066473).

REFERENCES

[1] Double robotics. https://www.doublerobotics.com. Accessed: April 4, 2019.
[2] Theta 360. https://theta360.com/en/about/theta/s.html. Accessed: April 4, 2019.
[3] F. Aurenhammer. Voronoi diagrams-a survey of a fundamental geometric data structure. *ACM Computing Surveys (CSUR)*, 1991.
[4] P. E. Hart et al. A formal basis for the heuristic determination of minimum cost paths. *IEEE Transactions on Systems Science and Cybernetics*, 4(2), 1968.
[5] I. Hwang et al. Talkbetter: Family-driven mobile intervention care for children with language delay. In *Proc. CSCW'14*. ACM, 2014.
[6] B. Kang et al. Zaturi: We put together the 25th hour for you. create a book for your baby. In *Proc. CSCW'17*. ACM, 2017.
[7] B. Kang et al. My being to your place, your being to my place: Co-present robotic avatars create illusion of living together. In *Proc. MobiSys'18*. ACM, 2018.
[8] J. Redmon and A. Farhadi. Yolov3: An incremental improvement. *arXiv preprint arXiv:1804.02767*, 2018.