Toward Personalized Tour-Guide Robot: Adaptive Content Planner based on Visitor's Engagement

Yanran Lin
Computer Science and Engineering
University of Michigan
Ann Arbor, MI, USA

Wonse Jo
Robotics Department
University of Michigan
Ann Arbor, MI, USA

Arsha Ali
Robotics Department
University of Michigan
Ann Arbor, MI, USA

Lionel P. Robert Jr.
School of Information
University of Michigan
Ann Arbor, MI, USA

Dawn M. Tilbury
Robotics Department
University of Michigan
Ann Arbor, MI, USA

ABSTRACT

In the evolving landscape of human-robot interactions, tour-guide robots are increasingly being integrated into various settings. However, the existing paradigm of these robots relies heavily on pre-recorded content, which limits effective engagement with visitors. We propose to address this issue of visitor engagement by transforming tour-guide robots into dynamic, adaptable companions that cater to individual visitor needs and preferences. Our primary objective is to enhance visitor engagement during tours through a robotic system capable of assessing and reacting to visitor preference and engagement. Leveraging this data, the system can calibrate and adapt the tour-guide robot's content in real-time to meet individual visitor preferences. Through this research, we aim to enhance the tour-guide robots' impact in delivering engaging and personalized visitor experiences by providing an adaptive tour-guide robot solution that can learn from humans' preferences and adapt its behaviors by itself.

CCS CONCEPTS

• **Human-centered computing** → **Scenario-based design**; *Systems and tools for interaction design.*

KEYWORDS

human-robot interaction, tour-guide robot, service robot, engagement perception, engagement generation.

ACM Reference Format:
Yanran Lin, Wonse Jo, Arsha Ali, Lionel P. Robert Jr., and Dawn M. Tilbury. 2024. Toward Personalized Tour-Guide Robot: Adaptive Content Planner based on Visitor's Engagement. In *Companion of the 2024 ACM/IEEE International Conference on Human-Robot Interaction (HRI '24 Companion), March 11–14, 2024, Boulder, CO, USA.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3610978.3640731

HRI '24, March 11–14, 2024, Boulder, CO, USA

ACM ISBN 979-8-4007-0323-2/24/03…$15.00
https://doi.org/10.1145/3610978.3640731

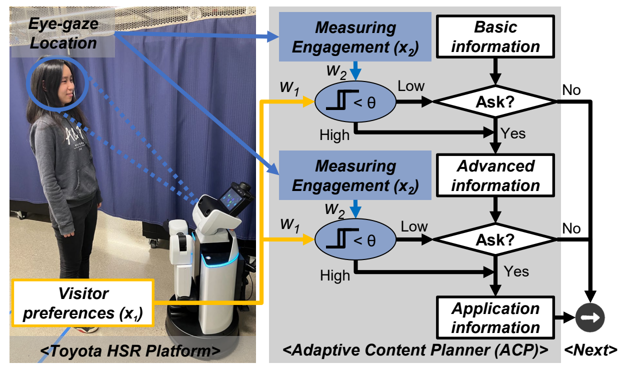


Figure 1: An overview of the proposed adaptive content planner (ACP) based on visitor engagement for a personalized tour-guide robot system.

1 INTRODUCTION

Tour-guide robots have been under development for more than two decades. One of the early examples was introduced in 1998 for museum tours [8]. Since then, several tour robot systems have been introduced, such as MINERVA [26], Robotinho [11], Lindsey [9], and GidaBot [19]. The main objective of these tour-guide robots is to guide visitors through complex environments, provide information, and enhance the overall visitor experience [24]. The advancements in artificial intelligence have enabled more recent developments in this field to offer a more personalized experience. Contemporary tour-guide robots are increasingly capable of adapting their information delivery based on visitor responses instead of delivering static information to the visitors, which limits the educational and informational impact of these robots [6]. A typical example is that Boston Dynamics turned its robot dog into a talking tour guide with large language models (LLM). They trained its four-legged bot to answer questions and generate responses about its facilities, which utilized humans' spoken language to adapt robots' behavior. However, a real-time interaction primarily relying on human voice input might not be feasible during many tours, where the visitors prefer to listen to the guide rather than keep talking. A self-learning system that adapts to the user's preferences is required[25]. Therefore, there still exist limitations to tour-guide robots' understanding of visitors' engagement and the reactions to this understanding.

To address this limitation, various types of interaction have been investigated on the tour-guide robots, such as spatial interaction [14], non-verbal interaction [3, 11, 15], and verbal interaction [21]. Spatial interaction refers to the robot and human sharing and reasoning about the same physical space. Non-verbal interaction emphasizes the interaction methods without speaking. To address the limitation mentioned above, we focus on the verbal interaction of the tour guide robot. Moreover, another intriguing avenue of the tour-guide robot field is the integration of learning technology [2, 9]. Although learning technology has been applied in guide robot design [23], it is still uncertain which is the best way to leverage this kind of technology to personalize the robot's behavior [20]. By employing machine learning and artificial intelligence techniques, the tour robots should have the potential to continually refine their interaction strategies, tailoring each tour to the unique characteristics of the audience. Therefore, we conclude that there exists a literature gap about how to combine learning technology and human-robot interaction methods in tour-guide robots for highly personalized and engaging tours.

To bridge this gap, we propose to develop a personalized robot tour framework to dynamically adjust the guide script content provided by the tour-guide robot based on measured visitor engagement. The architecture of this robot system is shown in Fig. 1. The definitions of variables in this system are provided in section III.

2 RELATED WORKS

2.1 Defining of Engagement

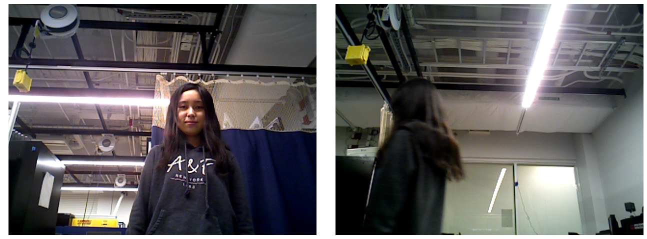
Engagement with an agent is typically referred to as social engagement including an affective component, such as emotion, workload, and entertainment. One way in which it is being used is to capture the inner state of a participant and the value they attribute to the interaction [16]. Youssef et al. [4] used several levels of engagement annotation including head rotation and eye gaze. In addition to the affective component, humans' attention and cognitive engagement matter in the engagement annotation as well. Biancardi et al. [5] combine affective- with attention- and cognitive engagement components in their detection model. Based on these statements about the definition of engagement, we define engagement as the attention and value the visitors attribute to the tour.

2.2 Measuring Engagement

Prior research has explored various engagement measurement methods in human-robot interactions. Post-interaction questionnaires were used in [22] to evaluate the engagement level of the humans. Pattar et al. [17] developed a module integrated with facial recognition, and speech recognition to measure engagement and enable personalized interaction. Some researchers took advantage of behavioral observations to measure the engagement level. Duchetto et al. [10] developed an engagement detector to measure visitor's engagement from their eye features captured from a robot's ego camera.

2.3 Calibrating Engagement

Studies have shown that adaptive robot behavior can enhance engagement, including speech patterns and communication strategies. Sidner et al. [22] improved visitor engagement by designing a robot to mimic human conversational gaze behavior. Oerte et al. [16] found that, in robot-user interaction, users are sensitive to the robot's gaze and gestures. Therefore, there are also some research works exploring aspects of gesture. Kanda et al. [13] proposed a robot architecture for implementing a large number of behaviors to entice people to relate to it interpersonally. Breazeal et al. [7] presented a biologically inspired framework for emotive communication and interaction between robots and humans with robot gesture designs. Velentza et al. [27] also found that the cheerful robot with more funny facial expressions and smiles could improve the visitors' engagement. Apart from the facial expressions, they manipulated the robots in terms of their voice and storytelling style. More cheerful voices and storytelling styles have a positive impact on the visitors' engagement as well.



(a) High engagement level **(b) Low engagement level**

Figure 2: Examples of the engagement detector results: (a) High engagement with the robot when visitors look at the tour-robot platform or the exhibit being discussed, (b) Low engagement with the robot when visitors pay no attention to the exhibit being discussed.

Based on our literature review, although there were some research works exploring engagement measures and enhancing engagement in human-robot interaction, there still exists a literature gap about how to enhance the engagement level in real-time based on a comprehensive assessment of the visitors' preference and engagement with a tour-guide robot.

3 ADAPTIVE CONTENT PLANNER BASED ON VISITOR'S ENGAGEMENT

Our proposed solution is an adaptive content planner (ACP) that is specifically designed to enhance visitor engagement during tours led by tour-guide robots. The ACP takes into account visitors' initial preferences for short or long tour that indicates their interest in this tour, and real-time engagement levels to estimate their interest in the content being presented. Based on the visitors' interest and real-time engagement, the ACP adjusts the content and the length of the guide scripts, helping the visitors remain interested and engaged throughout the robot tour. By doing so, we aim to provide a more personalized and enjoyable experience for visitors.

3.1 Measuring Preference and Engagement

As shown in Fig. 1, the ACP utilizes two inputs, initial preference (x_1) and engagement (x_2), to adjust the features of guide scripts during the interaction. The robot asks each visitor for their preferences for a short or long tour at the beginning of the tour. The visitor's response is recognized by speech recognition [29] and converted into the initial preference input x_1 for the ACP. For example, if the visitor chooses the long tour, a higher value (0.6) will be assigned to x_1 because they expect to spend more time on this tour, which indicates a relatively higher interest and value they attribute to this tour.

Gaze is an important cue involved in assessing engagement in human-robot interaction [1], which shows humans' attention to the robot. To measure real-time engagement from visitors, the robot observes the behaviors of the visitor with its camera and estimates the engagement level. We apply an existing deep learning-based engagement detector to estimate engagement levels during the tour-robot's explanations for each spot in the tour[10]. The engagement detector is built using the TOur GUide RObot (TOGURO) dataset, a collection of videos recorded at the museum, and annotated by humans. The engagement level is highest when the visitor is looking directly at the robot's head and behind the robot, medium when they are looking around within the robot camera view, and lowest when the visitor is not within the robot's camera view. According to their results in [10], Spearman's Correlation ρ of predictions with the ground-truth values annotated by three annotators is 0.758 for single-party interaction. Our tour-guide robot measures real-time visitor engagement using their engagement detector while discussing the exhibit at each spot. The average of the engagement values is considered as the engagement input x_2 for the ACP. The results of the engagement detector software tool with high and low engagement values are presented in Fig. 2.

3.2 Adjusting Contents Based on Visitors' Preference and Engagement

The ACP algorithm determines whether to ask the visitor to continue the explanation for the current spot or move to the next spot based on their preference and engagement. To make this decision, we used a linear threshold algorithm as shown in Eq. 1. It takes into account two inputs, namely the visitor's initial preference (x_1) and the measured engagement during the explanation (x_2).

$$f(x) = w_1 x_1 + w_2 x_2 \tag{1}$$

As shown in Fig. 1, the algorithm assigns a weight (w_1 and w_2) to each input respectively and compares the result to a threshold value (θ = 0.5) to determine whether to ask the visitor to continue or not. The output of the algorithm is a request for the visitor to continue or end the interaction.

The reason why we set the weights w_1 and w_2 is the uncertainty of the relative importance of input x_1 and x_2 in assessing engagement. The initial preference x_1 indicates the guide robot system is adaptable (e.g., the users can choose the exercises by themselves). In contrast, the real-time engagement level input x_2 indicates our system is adaptive (e.g., the system chooses exercises for the users by itself). Previous works show that users evaluate the adaptive robots as more competent, warm, and report a higher alliance[20]. To optimize both the engagement assessment and user experience, we set the weights w_1 and w_2 to make a reasonable balance between these two kinds of inputs. At this stage, the numerical values of x_1, w_1, w_2 and θ are set by iterative experiments. Several values of these variables and parameters were tested to optimize the adaptive interaction.

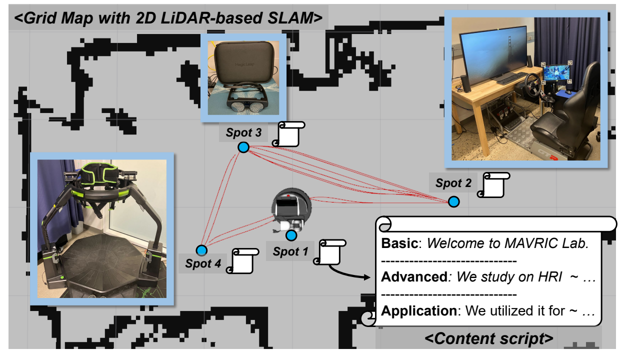


Figure 3: The experiment environment and tour spots. White paper icon represents an example of the guide scripts including basic, advanced, and application introductions.

4 PRELIMINARY EXPERIMENT AND RESULTS

We designed a robot-tour scenario at a lab space to validate the functionality of the proposed ACP. The mission of the tour guide robot is to introduce our lab research based on lab facilities. There is a supplementary video at https://youtu.be/WvgnWXWxLwk that explains the details of this robot-guide tour experiment.

4.1 Tour-guide Robot System

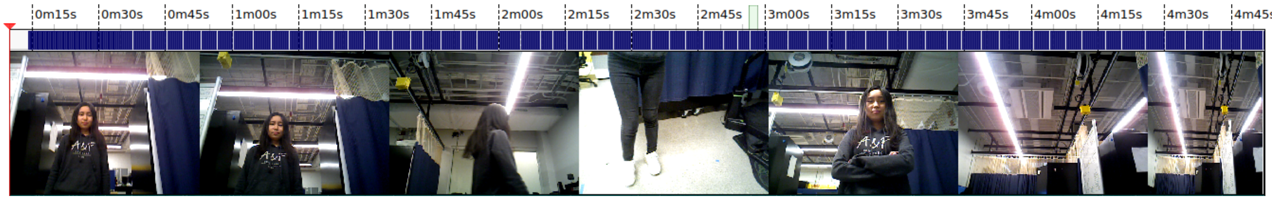
The tour-robot platform used in this experiment is Toyota Human Support Robot (HSR) [28], a mobile manipulator platform capable of navigating in the indoor environment. The HSR features various functions necessary for tour guide, such as robot speaking and speech detection. To easily integrate HSR functions with our ACP algorithm, we utilized the Robot Operating System (ROS) [18].

For the robot navigation through predefined spots, we built a grid map with Simultaneous Localisation and Mapping (SLAM) using gmapping package [12], and validated the mobility of the tour-robot navigation to avoid obstacles and dynamic visitors during the tour guide. Fig 3 shows the lab environment and grid map used in the tour-robot navigation, in which the red dash line indicates the robot's trajectories and the blue dot indicates the tour spots from Spot 1 to Spot 4.

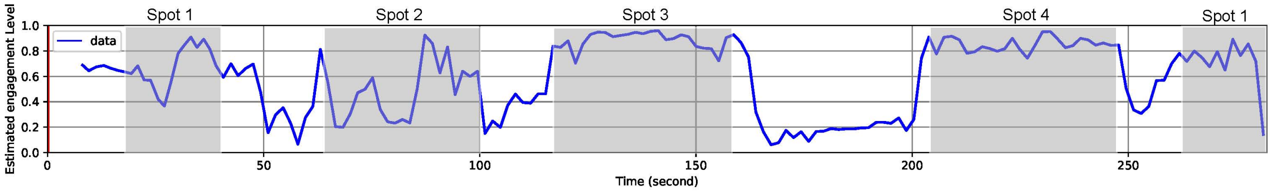
4.2 Experiment Scenario

Spot 1 to Spot 4 are the predefined spots for robot navigation, as illustrated in Fig. 3. Spot 1 is the beginning and end spot of the tour. Spot 2 to Spot 4 are the locations of the lab facilities. Each spot has three levels of introduction content in the corresponding guide script: basic, advanced, and application. The basic script contains basic information for the research facility, such as name, specification, and research objectives. The advanced script covers more technical details about the facility and research work. And, the application script focuses on the practical applications and real-world implications of the research work.

In our experiment scenario, at Spot 1, the tour robot first gave a brief introduction of our lab, then asked the visitor's preference on the tour mode (i.e., short or long tour) via robot voices (*"Would you prefer short tour? Please answer after Beep sound. Beep!"*). After *"Beep"* sound, the speech recognition algorithm of the HSR platform



(a) HSR Robot's camera view



(b) Estimated engagement level during experiment

Figure 4: One of the preliminary experiment's results: The top image displays the HSR's camera views, and the bottom image shows the estimated engagement level during the robot tour. The gray box indicates the tour-guide robot's explanation of the content from Spot 1 to Spot 4.

detects the visitor's answer which becomes input x_1 in the ACP. Then, the robot moved to the next spot to continue the robot tour. As the tour robot leads visitors to each spot from Spot 2 to 4, it firstly provides the basic information. While explaining the basic content script, the robot measures the visitors' engagement level values with a 10 Hz sampling rate from the robot camera views mounted on the HSR head. When the basic explanation finishes, the mean value of these engagement level values is considered as input x_2. Based on the APC result, the robot decides whether to continue explaining more details of the current spot. This process repeats during the robot-guide tour.

4.3 Alpha Test and Results

We conducted alpha tests to verify the functioning of the tour robot system and ACP algorithm. The robot led a tour for one visitor each time. During the tests, participants were instructed to pay attention to the robot guide or not at different spots. We presented one of the test results in Fig. 4, in which Fig. 4b displays the engagement levels during the test and Fig. 4a shows the camera view from the HSR platform.

During the robot explanation at Spot 2, the participant showed disinterest in the tour by looking around. At Spot 3, the participant showed interest in the robot by closely eye-contacting with the robot and watching the lab facility. As a result, the robot should ask the participants if they want to learn more about the current spot at Spot 2. On the other hand, at Spot 3, the robot should continue with the advanced script to cover more details without asking the participant. Under this circumstance, each tour took around 5 min with around 300 words spoken. It takes around 15 sec for the robot to go from one spot to another. Most of the alpha tests showed the same results as we expected, which validated that the robot system can adjust the content based on the visitor's engagement. However, sometimes the system did give some unexpected response. For example, the system might ask the visitor if they would like to continue with an advanced script even if the visitor felt engaged at that time.

5 CONCLUSIONS AND FUTURE WORK

We proposed a new interactive system for tour-guide robots that is capable of measuring visitors' engagement and using this data to customize the content provided by the tour-guide robot according to individual needs. To implement this system, we used the Toyota HSR platform to navigate predefined spots in our experimental environment (i.e., the lab space). We demonstrated the system's performance and capabilities in playing a tour-guide role, such as speech recognition, measuring visitor engagement from the robot's camera view, and explaining the spot with three levels of guide scripts based on the visitor's preferences and real-time engagement level.

However, there are some limitations in our proposed primary experiment. Firstly, we did not deeply consider optimizing the values of variables and parameters in the ACP algorithm. Therefore, it will be necessary to adjust and standardize these values, including the conversion from visitor's preference to input x_1, the weights of preference and engagement values (w_1 and w_2), and the threshold of the engagement level. Secondly, we did not validate the proposed ACP algorithm with multiple human visitors. As we consider real-world scenarios, it is highly likely that multiple visitors will be present in the environment simultaneously. Lastly, we assumed the visitor was not interested in the tour when the visitor kept looking around or leaving the robot's view in the alpha tests. But this hypothesis might not be always valid.

In the future, we plan to overcome the aforementioned limitations and conduct a user experiment to thoroughly validate the proposed ACP algorithm. Furthermore, we intend to investigate the impact of dynamic content adaptation on visitor engagement and test the proposed system via user experiments.

REFERENCES

[1] Henny Admoni and Brian Scassellati. 2017. Social eye gaze in human-robot interaction: a review. *Journal of Human-Robot Interaction* 6, 1 (2017), 25–63.

[2] Víctor Alvarez-Santos, Adrián Canedo-Rodriguez, Roberto Iglesias, Xosé M Pardo, Carlos V Regueiro, and M Fernandez-Delgado. 2015. Route learning and reproduction in a tour-guide robot. *Robotics and Autonomous Systems* 63 (2015), 206–213.

[3] V. Alvarez-Santos, R. Iglesias, X.M. Pardo, C.V. Regueiro, and A. Canedo-Rodriguez. 2014. Gesture-based interaction with voice feedback for a tour-guide robot. *Journal of Visual Communication and Image Representation* 25, 2 (2014), 499–509. https://doi.org/10.1016/j.jvcir.2013.03.017

[4] Atef Ben-Youssef, Giovanna Varni, Slim Essid, and Chloé Clavel. 2019. On-the-fly detection of user engagement decrease in spontaneous human–robot interaction using recurrent and deep neural networks. *International Journal of Social Robotics* 11, 5 (2019), 815–828. https://doi.org/10.1007/s12369-019-00591-2

[5] Beatrice Biancardi, Maurizio Mancini, Paul Lerner, and Catherine Pelachaud. 2019. Managing an agent's self-presentational strategies during an interaction. *Frontiers in Robotics and AI* 6 (2019). https://doi.org/10.3389/frobt.2019.00093

[6] Hans-Joachim Böhme, Torsten Wilhelm, Jürgen Key, Carsten Schauer, Christof Schröter, Horst-Michael Groß, and Torsten Hempel. 2003. An approach to multi-modal human–machine interaction for intelligent service robots. *Robotics and Autonomous Systems* 44, 1 (2003), 83–96.

[7] Cynthia Breazeal. 2001. Affective interaction between humans and robots. In *European conference on artificial life*. Springer, 582–591.

[8] Wolfram Burgard, Armin B Cremers, Dieter Fox, Dirk Hähnel, Gerhard Lakemeyer, Dirk Schulz, Walter Steiner, and Sebastian Thrun. 1998. The Interactive Museum Tour-Guide Robot.. In *Aaai/iaai*. 11–18.

[9] Francesco Del Duchetto, Paul Baxter, and Marc Hanheide. 2019. Lindsey the tour guide robot-usage patterns in a museum long-term deployment. In *2019 28th IEEE international conference on robot and human interactive communication (RO-MAN)*. IEEE, 1–8.

[10] Francesco Del Duchetto, Paul Baxter, and Marc Hanheide. 2020. Are you still with me? continuous engagement assessment from a robot's point of view. *Frontiers in Robotics and AI* 7 (2020). https://doi.org/10.3389/frobt.2020.00116

[11] Felix Faber, Maren Bennewitz, Clemens Eppner, Attila Gorog, Christoph Gonsior, Dominik Joho, Michael Schreiber, and Sven Behnke. 2009. The humanoid museum tour guide Robotinho. In *RO-MAN 2009-The 18th IEEE International symposium on robot and human interactive communication*. IEEE, 891–896.

[12] Giorgio Grisetti, Cyrill Stachniss, and Wolfram Burgard. 2019. Gmapping, a highly efficient Rao-Blackwellized particle filer to learn grid maps from laser range data. *URL: http://openslam. org/gmapping. html [cited 06-08-2016]* (2019).

[13] Takayuki Kanda, Hiroshi Ishiguro, Michita Imai, Tetsuo Ono, and Kenji Mase. 2002. A constructive approach for developing interactive humanoid robots. In *IEEE/RSJ international conference on intelligent robots and systems*, Vol. 2. IEEE, 1265–1270.

[14] Tribhi Kathuria, Yifan Xu, Theodor Chakhachiro, X Jessie Yang, and Maani Ghaffari. 2022. Providers-Clients-Robots: Framework for spatial-semantic planning for shared understanding in human-robot interaction. In *2022 31st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 1099–1106.

[15] Lilia Moshkina, Susan Trickett, and J Gregory Trafton. 2014. Social engagement in public places: a tale of one robot. In *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction*. 382–389.

[16] Catharine Oertel, Ginevra Castellano, Mohamed Chetouani, Jauwairia Nasir, Mohammad Obaid, Catherine Pelachaud, and Christopher Peters. 2020. Engagement in human-agent interaction: An overview. *Frontiers in Robotics and AI* 7 (2020). https://doi.org/10.3389/frobt.2020.00092

[17] Suraj Prakash Pattar, Enrique Coronado, Liz Rincon Ardila, and Gentiane Venture. 2019. Intention and Engagement Recognition for Personalized Human-Robot Interaction, an integrated and Deep Learning approach. In *2019 IEEE 4th International Conference on Advanced Robotics and Mechatronics (ICARM)*. 93–98. https://doi.org/10.1109/ICARM.2019.8834226

[18] Morgan Quigley, Ken Conley, Brian Gerkey, Josh Faust, Tully Foote, Jeremy Leibs, Rob Wheeler, Andrew Y Ng, et al. 2009. ROS: an open-source Robot Operating System. In *ICRA workshop on open source software*, Vol. 3. Kobe, Japan, 5.

[19] Igor Rodriguez, Unai Zabala, Pedro A Marín-Reyes, Ekaitz Jauregi, Javier Lorenzo-Navarro, Elena Lazkano, and Modesto Castrillón-Santana. 2020. Personal guides: heterogeneous robots sharing personal tours in multi-floor environments. *Sensors* 20, 9 (2020), 2480.

[20] Sebastian Schneider and Franz Kummert. 2021. Comparing robot and human guided personalization: adaptive exercise robots are perceived as more competent and trustworthy. *International Journal of Social Robotics* 13, 2 (2021), 169–185.

[21] Soobin Seo. 2022. When Female (Male) Robot Is Talking To Me: Effect of service robots' gender and anthropomorphism on customer satisfaction. *International Journal of Hospitality Management* 102 (2022), 103166.

[22] Candace L Sidner, Cory D Kidd, Christopher Lee, and Neal Lesh. 2004. Where to look: a study of human-robot engagement. In *Proceedings of the 9th international conference on Intelligent user interfaces*. 78–84.

[23] Mohsen Soori, Behrooz Arezoo, and Roza Dastres. 2023. Artificial intelligence, machine learning and deep learning in advanced robotics, A review. *Cognitive Robotics* (2023).

[24] Ronny Stricker, Steffen Müller, Erik Einhorn, Christof Schröter, Michael Volkhardt, Klaus Debes, and Horst-Michael Gross. 2012. Konrad and Suse, two robots guiding visitors in a university building. *Informatik aktuell* (2012), 49–58. https://doi.org/10.1007/978-3-642-32217-4_6

[25] Karthick Thiyagarajan, Sarath Kodagoda, Mark Luu, Taliah Duggan-Harper, Doug Ritchie, Kelly Prentice, and Jodi Martin. 2022. Intelligent Guide Robots for People who are Blind or have Low Vision: A Review. *Vision Rehabilitation International* 13, 1 (2022), 1–15. https://doi.org/doi:10.2478/vri-2022-0003

[26] Sebastian Thrun, Maren Bennewitz, Wolfram Burgard, Armin B Cremers, Frank Dellaert, Dieter Fox, Dirk Hahnel, Charles Rosenberg, Nicholas Roy, Jamieson Schulte, et al. 1999. MINERVA: A second-generation museum tour-guide robot. In *Proceedings 1999 IEEE International Conference on Robotics and Automation (Cat. No. 99CH36288C)*, Vol. 3. IEEE.

[27] Anna-Maria Velentza, Dietmar Heinke, and Jeremy Wyatt. 2019. Human interaction and improving knowledge through collaborative tour guide robots. In *2019 28th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 1–7.

[28] Takashi Yamamoto, Koji Terada, Akiyoshi Ochiai, Fuminori Saito, Yoshiaki Asahara, and Kazuto Murase. 2018. Development of the research platform of a domestic mobile manipulator utilized for international competition and field test. In *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 7675–7682.

[29] Anthony Zhang. 2017. Speech Recognition (Version 3.8). https://github.com/Uberi/speech_recognition#readme