

Human-Robot Interaction

Subjects: Robotics

Contributor: Sherif Said

As in human–human interaction, several modalities can be used at once in human-robot interaction in social contexts. Vision, eye gaze, verbal dialogue, touch, and gestures are examples of modalities that can be used herein. In a social context, the intelligence that a robot display depends on the modalities it uses and each modality can have specific importance and effect on the human side of the interaction which translates into the degree of trust that the robot has. Moreover, the acceptance of robots in social interaction depends on their ability to express emotions and they require a proper design of emotional expressions to improve their likability and believability as multimodal interaction can enhance the engagement

Keywords: social robotics ; artificial intelligence ; human-robot interaction

1. Modalities of Human-Robot Interaction

1.1. Vision Systems in Robots

Visual perception provides what was suggested to be the most important information to robots, allowing them to achieve successful interaction with human partners ^[1]. This information can be used in a variety of tasks, such as navigation, obstacle avoidance, detection, understanding, and manipulation of objects, and assigning meanings to a visual configuration of a scene ^{[2][3]}. More specifically, the vision has been used for the estimation of the 3D position and orientation of a user in an environment ^[4], the estimation of distances between a robot and users ^[5], tracking human targets and obtaining their poses ^[6], understanding human behavior aiming to contribute to the cohabitation between assistive robots and humans ^[7]. Similarly, the vision has been used in a variety of other applications, such as recognizing patterns and figures in exercises in a teaching assistance context in a high school ^[8], detecting and classifying waste material as a child would do ^[9], and detecting people entering a building for a possible interaction ^[10]. Moreover, the vision has been used in ^[11] for medication sorting, taking into account pill types and numbers, in ^[12] for sign recognition in a sign tutoring task with deaf or hard of hearing children, and in ^[13] as part of a platform used for cognitive stimulation in elderly users with mild cognitive impairments.

1.2. Conversational Systems in Robots

Although some applications of social robotics involve robots taking vocal commands without generating a vocal reply ^[14], interactions can be made richer when the robot can engage in conversations. A typical social robot with autonomous conversation ability must have the capacity to acquire sound signals, process them to recognize the speech, recognize the whole sequence of words pronounced by the human interlocutor, formulate an appropriate reply, and synthesize the sound signal corresponding to the reply, then emit this signal using a loudspeaker. The core component of this ability is the recognition of word sequences and the generation of reply sequences ^[15]. This can rely on a learning stage where the system acquires the experience of answering word sequences by observing a certain number of conversations that are mainly between humans. Techniques used in this area involve word and character embeddings, and learning through recurrent neural network (RNN) architectures, long short-term memory networks (LSTM), and gated recurrent units (GRU) ^{[15][16]}. It is to note that not all social robotic systems with conversational capacities have the same levels of complexity as some use limited vocabularies in their verbal dialogues. Herein, Conversation scenarios were seen in ^[17], verbal dialogue in ^[18], dialogues between children and a robot in ^[19], and some word utterances in ^[9]. A review of conversational systems usages in psychiatry was made in ^[20]. It covered different aspects such as therapy bots, avatars, and intelligent animal-like robots. Additionally, an algorithm for dialogue management has been proposed in ^[21] for social robots and conversational agents. It is aimed at ensuring a rich and interesting conversation with users. Furthermore, robot rejection of human commands has been addressed in ^[22] with aspects such as how rejections can be phrased by the robot. GPT-3 ^[23] has emerged as a language model with potential applications in conversational systems and social robotics ^[24]. However, in several conversational systems, problems have been reported, such as hallucinations ^{[25][26]}, response blandness, and incoherence ^[15]. The research work presented in ^[27] aimed at improving the conversational capabilities of

a social robot by reducing the possibility of problems as described above, and improving the human-robot interaction with an expressive face. It intended to have a 3-D printed animatronic robotics head with an eye mechanism, a jaw mechanism, and a head mechanism. The three mechanisms are designed to be driven by servo motors to actuate the head synchronously with the audio output. The robotics head design is optimized to fit microphones, cameras, and speakers. The robotics head is envisioned to meet students and visitors in a university. To ensure the appropriateness of the interactions, several stages will be included in the control framework of the robot and a database of human-human conversations will be built upon for the machine learning of the system. This database will be built in the aim of training the conversational system in contexts similar to its contexts of usage. This will increase the adequacy of the conversational system's parameters with respect to the tasks it is required to do, and increase the coherence and consistency of the utterances it produces. For that, the recorded data will comply with the following specifications:

- Context of the interactions: Visitors approach the receptionist and engage in conversations in English. Both questions and answers will be included in the database.
- Audio recordings of the conversations: a text by speech recognition modules is used to transcript the conversation.
- Video recordings of the interaction, showing the face and upper body of the receptionist, with a quality of images usable by body posture recognition systems.
- The collected data will be used to progressively train the system. Each conversation will be labeled with the corresponding date, time and interaction parties.
- Participants will be asked to be free to ask questions they may have to inquire about the center in English, without having any other constraint or any specific text to pronounce.

1.3. Expressions and Gestures

Aside from the ability to process and generate sequences of words, a social robot requires more capacities to increase engagement and realism in the interaction with a human. This can be done through speech-accompanying gestures and facial expressions. Indeed, facial expression has an important role in communication between humans because it is rich in information, together with gestures and sound [28][29][30]. This issue has been studied in psychology, and research indicates that there are six main emotions associated with distinctive facial expressions [31]. At Columbia University [32], scientists and engineers developed a robot that can raise eyebrows, smile, and have forehead wrinkles similar to humans. This robot can express the face more accurately compared to the rest of the robots. This robot, called Eva, can mimic head movements and facial expressions. In this robot, 25 muscles are used, and 12 of them are dedicated specifically to the face. These muscles can produce facial skin excitations of up to 15 mm. In other works, different examples can be found for applications of gestures and expressions in social robotics. For instance, gestures have been combined with verbal dialogue and screen display in [18] for health data acquisition in hospitals with Pepper. In [33], a robot with the ability to display facial expressions was used in studies related to storytelling robots. These studies focused on the roles of the emotional facial display, contextual head movements, and voice acting. In [29], a framework for generating robot behaviors using speech, gestures, and facial expressions was proposed, to improve the expressiveness of a robot in interaction with humans. of interaction with human users.

2. Metrics of Human Perception and Acceptability

The usage of social robots in the different environments and contexts presented above is subjected to their acceptability by humans as partners in the interaction. Indeed, to be accepted in social contexts, robots need to show degrees of intelligence, morphology, or usefulness that can be judged positively by users, not to mention cultural influences on expectations towards and responses to social robots [34]. The study published in 2021 in [35] focused on the perception that humans have of the cognitive and affective abilities of robots and began with the hypothesis that this perception varied in accordance with the degree of human-likeness that robots have. However, the results obtained with students on four robots used in the study did not prove this hypothesis. A study made in 2005 in [36] showed the acceptability of persons for robots as companions in the home, more as assistants, machines, or servants than as a friend. More recently, the literature review and study made in [37] mentions anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety as five key concepts in human-robot interaction. The study also emphasized the importance of being aware of human perception and cognition measures developed by psychologists for engineers developing robots. Additionally, according to the tasks expected from the robots, different measures of performance can be made, such as true recognition measures in speech recognition tasks. But a robot can have a high performance in a specific task, without having a positive impact on its social context. Therefore, the performances of robots in social usage are in many cases

measured through evaluations made by humans using questionnaires and metrics calculated based on them. The outcomes of such evaluations are affected by the subjectivity of the persons participating in them and their numbers. Herein, certain metrics/measures can be mentioned as follows:

- in [38], a robotic platform was equipped with the capacity to perform the two tasks of group interaction, where it had to maintain an appropriate position and orientation in a group, and the person following. The human evaluation began with a briefing of 15 subjects about the purpose of each task, followed by a calibration step where the subjects were shown human-level performance in each task, followed by interaction with the robotic platform for each task. Then, the subjects were asked to rate the social performance of the platform with a number from 1 to 10 where 10 was human-level performance. The authors suggested increasing the number of subjects and a more detailed questionnaire to be necessary for reaching definitive conclusions.
- the “Godspeed” series of questionnaires has been proposed in [37] to help creators of robots in the robot development process. Five questionnaires using 5-point scales address the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. For example, in the anthropomorphism questionnaire (Godspeed I), participants are asked to rate their impressions of the robot with an integer from fake (1) to natural (5), and from machine-like (1) to human-like (5), and from artificial (1) to lifelike (5). Also in the animacy questionnaire (Godspeed II), participants can rate the robot for example from dead (1) to alive (5), from stagnant (1) to lively (5), and from inert (1) to interactive (5). The authors in [37] report cultural backgrounds, prior experiences with robots, and personality to be among the factors affecting the measurements made in such questionnaires. Furthermore, the perceptions of humans are unstable as their expectations and knowledge change with the increase of their experiences with robots. This means, for the authors in [37], that repeating the same experiment after a long duration of time would yield different results.
- in the context of elderly care and assistance, the Almere model was proposed in [39] as an adaptation and theoretical extension of the Unified Theory of Acceptance and Use of Technology (UTAUT) questionnaire [40]. Questionnaire items in the Almere model were adapted from the UTAUT questionnaire to fit the context of assistive robot technology and address elderly users in a care home. Different constructs are adopted and defined and questionnaires related to them, respectively. This resulted in constructs such as the users’ attitude towards the technology their intention to use it, their perceived enjoyment, perceived ease of use, perceived sociability and usefulness, social influence and presence, and trust. Experiments made on the model consisted of a data collection instrument with different questionnaire items on a 5-point Likert-type scale ranging from 1 to 5, corresponding to statements ranging from “totally disagree” to “totally agree”, respectively.

Other metrics and approaches for the evaluation of the engagement in the interaction between humans and robots have been proposed. The work presented in [41] proposes metrics that can be easily retrieved from off-the-shelf sensors, by static and dynamic analysis of body posture, head movements and gaze of the human interaction partner.

The work made in [37] revealed two important points related to the assessment of human-robot interaction: the need for a standardized measurement tool and the effects of user background and time on the measurements. The authors also invited psychologists to contribute to the development of the questionnaires. These issues can have implications for social robotics studies that should be addressed to improve assessment quality and results and advance robotic system designs and tasks accurately. More recently, the work shown in [42] proposed a standardized process for choosing and using scales and questionnaires used in human-robot interaction. For instance, the authors in [42] specified that a scale cannot be trusted in a certain study if not already validated in a similar study and that scales can be unfit or have limitations concerning a specific study. In such a case, they should be modified and re-validated.

References

1. Kang, S.H.; Han, J.H. Video Captioning Based on Both Egocentric and Exocentric Views of Robot Vision for Human-Robot Interaction. *Int. J. Soc. Robot.* 2021.
2. Kragic, D.; Vincze, M. Vision for Robotics. *Found. Trends Robot.* 2010, 1, 1–78.
3. Ronchi, M.R. Vision for Social Robots: Human Perception and Pose Estimation. Ph.D. Thesis, California Institute of Technology, Pasadena, CA, USA, 2020.
4. Garcia-Salguero, M.; Gonzalez-Jimenez, J.; Moreno, F.A. Human 3D Pose Estimation with a Tilting Camera for Social Mobile Robot Interaction. *Sensors* 2019, 19, 4943.

5. Pathi, S.K.; Kiselev, A.; Kristoffersson, A.; Repsilber, D.; Loutfi, A. A Novel Method for Estimating Distances from a Robot to Humans Using Egocentric RGB Camera. *Sensors* 2019, 19, 3142.
6. Gonzalez-Pacheco, V.; Ramey, A.; Alonso-Martin, F.; Castro-Gonzalez, A.; Salichs, M.A. Maggie: A Social Robot as a Gaming Platform. *Int. J. Soc. Robot.* 2011, 3, 371–381.
7. Kostavelis, I.; Vasileiadis, M.; Skartados, E.; Kargakos, A.; Giakoumis, D.; Bouganis, C.S.; Tzovaras, D. Understanding of Human Behavior with a Robotic Agent through Daily Activity Analysis. *Int. J. Soc. Robot.* 2019, 11, 437–462.
8. Reyes, G.E.B.; Lopez, E.; Ponce, P.; Mazon, N. Role Assignment Analysis of an Assistive Robotic Platform in a High School Mathematics Class, Through a Gamification and Usability Evaluation. *Int. J. Soc. Robot.* 2021, 13, 1063–1078.
9. Castellano, G.; De Carolis, B.; D'Errico, F.; Macchiarulo, N.; Rossano, V. PeppeRecycle: Improving Children's Attitude Toward Recycling by Playing with a Social Robot. *Int. J. Soc. Robot.* 2021, 13, 97–111.
10. Saad, E.; Broekens, J.; Neerincx, M.A.; Hindriks, K.V. Enthusiastic Robots Make Better Contact. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Macau, China, 3–8 November 2019.
11. Wilson, J.R.; Lee, N.Y.; Saechao, A.; Tickle-Degnen, L.; Scheutz, M. Supporting Human Autonomy in a Robot-Assisted Medication Sorting Task. *Int. J. Soc. Robot.* 2018, 10, 621–641.
12. Gurpinar, C.; Uluer, P.; Akalin, N.; Kose, H. Sign Recognition System for an Assistive Robot Sign Tutor for Children. *Int. J. Soc. Robot.* 2020, 2, 355–369.
13. Cosar, S.; Fernandez-Carmona, M.; Agrigoroaie, R.; Pages, J.; Ferland, F.; Zhao, F.; Yue, S.; Bellotto, N.; Tapus, A. ENRICHME: Perception and Interaction of an Assistive Robot for the Elderly at Home. *Int. J. Soc. Robot.* 2020, 12, 779–805.
14. Al-Abdullah, A.; Al-Ajmi, A.; Al-Mutairi, A.; Al-Mousa, N.; Al-Daihani, S.; Karar, A.S.; Alkork, S. Artificial Neural Network for Arabic Speech Recognition in Humanoid Robotic Systems. In *Proceedings of the 2019 3rd International Conference on Bio-engineering for Smart Technologies (BioSMART)*, Paris, France, 24–26 April 2019; pp. 1–4.
15. Gao, J.; Galley, M.; Li, L. *Neural Approaches to Conversational AI, Question Answering, 1052 Task-Oriented Dialogues and Social Chatbots*; Now Foundations and Trends: Hanover, MA, USA, 2019.
16. Dzakwan, G.; Purwarianti, A. Comparative Study of Topology and Feature Variants for Non-Task-Oriented Chatbot using Sequence to Sequence Learning. In *Proceedings of the 5th International Conference on Advanced Informatics: Concept Theory and Applications (ICAICTA)*, Krabi, Thailand, 14–17 August 2018.
17. Obayashi, K.; Kodate, N.; Masuyama, S. Assessing the Impact of an Original Soft Communicative Robot in a Nursing Home in Japan: Will Softness or Conversations Bring more Smiles to Older People? *Int. J. Soc. Robot.* 2022, 14, 645–656.
18. Van der Putte, D.; Boumans, R.; Neerincx, M.; Rikkert, M.O.; De Mul, M. A Social Robot for Autonomous Health Data Acquisition among Hospitalized Patients: An Exploratory Field Study. In *Proceedings of the 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, Daegu, Korea, 11–14 March 2019.
19. Ismail, L.I.; Hanapih, F.A.; Belpaeme, T.; Dambre, J.; Wyffels, F. Analysis of Attention in Child-Robot Interaction Among Children Diagnosed with Cognitive Impairment. *Int. J. Soc. Robot.* 2021, 13, 141–152.
20. Pham, K.T.; Nabizadeh, A.; Selek, S. Artificial Intelligence and Chatbots in Psychiatry. *Psychiatr. Q.* 2022, 93, 249–253.
21. Grassi, L.; Recchiuto, C.T.; Sgorbissa, A. Knowledge-Grounded Dialogue Flow Management for Social Robots and Conversational Agents. *Int. J. Soc. Robot.* 2022.
22. Briggs, G.; Williams, T.; Jackson, R.B.; Scheutz, M. Why and How Robots Should Say 'No'. *Int. J. Soc. Robot.* 2022, 14, 323–339.
23. Brown, T.B.; Mann, B.; Ryder, N.; Subbiah, M.; Kaplan, J.; Dhariwal, P.; Neelakantan, A.; Shyam, P.; Sastry, G.; Askell, A.; et al. Language Models are Few-Shot Learners. *Adv. Neural Inf. Process. Syst.* 2020, 33, 1877–1901.
24. Psychiatry.org—DSM. Available online: www.dsm5.org (accessed on 5 January 2022).
25. Shuster, K.; Poff, S.; Chen, M.; Kiela, D.; Weston, J. Retrieval Augmentation Reduces Hallucination in Conversation. *arXiv* 2021, arXiv:2104.07567.
26. Maynez, J.; Narayan, S.; Bohnet, B.; McDonald, R. On Faithfulness and Factuality in Abstractive Summarization. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*; Association for Computational Linguistics, Online, 5–10 July 2020; pp. 1906–1919.
27. Youssef, K.; Said, S.; Beyrouthy, T.; Alkork, S. A Social Robot with Conversational Capabilities for Visitor Reception: Design and Framework. In *Proceedings of the 2021 4th International Conference on Bio-Engineering for Smart Technologies (BioSMART)*, Paris/Créteil, France, 8–10 December 2021; pp. 1–4.

28. Hazourli, A.; Djeghri, A.; Salam, H.; Othmani Morgan, A. Multi-facial patches aggregation network for facial expression recognition and facial regions contributions to emotion display. *Multimed. Tools Appl.* 2021, 80, 13639–13662.
29. Aly, A.; Tapus, A. On Designing Expressive Robot Behavior: The Effect of Affective Cues on Interaction. *SN Comput. Sci.* 2020, 1, 314.
30. Boucenna, S.; Gaussier, P.; Andry, P.; Hafemeister, L. Imitation as a Communication Tool for Online Facial Expression Learning and Recognition. In *Proceedings of the 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Taipei, Taiwan, 18–22 October 2010; pp. 5323–5328.
31. Ashraf, A.B.; Lucey, S.; Cohn, J.F.; Chen, T.; Ambadar, Z.; Prkachin, K.M.; Solomon, P.E. The painful face—pain expression recognition using active appearance models. *Image Vis. Comput.* 2009, 27, 1788–1796.
32. Faraj, Z.; Selamat, M.; Morales, C.; Torres, P.; Hossain, M.; Chen, B.; Lipson, H. Facially expressive humanoid robotic face. *HardwareX* 2021, 9, e00117.
33. Striepe, H.; Donnermann, M.; Lein, M.; Lugin, B. Modeling and Evaluating Emotion, Contextual Head Movement and Voices for a Social Robot Storyteller. *Int. J. Soc. Robot.* 2021, 13, 441–457.
34. Lim, V.; Rooksby, M.; Cross, E.S. Social Robots on a Global Stage: Establishing a Role for Culture During Human–Robot Interaction. *Int. J. Soc. Robot.* 2021, 13, 1307–1333.
35. Fortunati, L.; Manganelli, A.M.; Hoflich, J.; Ferrin, G. Exploring the Perceptions of Cognitive and Affective Capabilities of Four, Real, Physical Robots with a Decreasing Degree of Morphological Human Likeness. *Int. J. Soc. Robot.* 2021.
36. Dautenhahn, K.; Woods, S.; Kaouri, C.; Walters, M.; Koay, K.; Werry, I. What is a robot companion-Friend, assistant or butler? In *Proceedings of the 2005 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Edmonton, AB, Canada, 2–6 August 2005; pp. 1192–1197.
37. Bartneck, C.; Kulic, D.; Croft, E.; Zoghbi, S. Measurement Instruments for the Anthropomorphism, Animacy, Likeability, Perceived Intelligence, and Perceived Safety of Robots. *Int. J. Soc. Robot.* 2009, 1, 71–81.
38. Shiarlis, K.; Messias, J.; Whiteson, S. Acquiring Social Interaction Behaviours for Telepresence Robots via Deep Learning from Demonstration. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Vancouver, BC, Canada, 24–28 September 2017.
39. Heerink, M.; Kroese, B.; Evers, V.; Wielinga, B. Assessing Acceptance of Assistive Social Agent Technology by Older Adults: The Almere Model. *Int. J. Soc. Robot.* 2010, 2, 361–375.
40. Venkatesh, V.; Morris, M.G.; Davis, G.B.; Davis, F.D. User Acceptance of Information Technology: Toward a Unified View. *Manag. Inf. Syst. Q.* 2003, 27, 425–478.
41. Anzalone, S.; Boucenna, S.; Ivaldi, S.; Chetouani, M. Evaluating the Engagement with Social Robots. *Int. J. Soc. Robot.* 2015, 7, 465–478.
42. Rueben, M.; Elprama, S.A.; Chrysostomou, D.; Jacobs, A. Introduction to (re)using questionnaires in human-robot interaction research. In *Human-Robot Interaction: Evaluation Methods and Their Standardization*; Jost, C., Le Pévédic, B., Belpaeme, T., Bethel, C., Chrysostomou, D., Crook, N., Grandgeorge, M., Mirnig, N., Eds.; Springer International Publishing: Cham, Switzerland, 2020; pp. 125–144.