International Journal on Recent Researches in Science, Engineering & Technology (IJRRSET)
A Journal Established in early 2000 as National journal and upgraded to International journal in 2013 and is in existence for the last 10 years. It is run by Retired Professors from NIT, Trichy. Journal Indexed in JIR, DIIF and SJIF.
Available online at: www.jrrset.com

ISSN (Print) : 2347-6729
ISSN (Online) : 2348-3105
Volume 8, Issue 1
January 2020.
JIR IF : 2.54
SJIF IF : 4.334
Cosmos: 5.395

Received on: 04.12.2019 Published on: 20.01.2020 Pages : 30-36

SOCIAL ROBOTS FOR PUBLIC INTERACTION

L. Ranganathan[a*], R. Nandhakumar [b], T.Ravichandra Babu[c]

[a,b] Department of Mechatronics Engineering, Agni College of Technology, Chennai 600130, Tamilnadu, India
[c]Department of Electronics and Communication Engineering, Agni College of Technology, Chennai 600130, Tamilnadu, India

[a] e-mail: ranganathan.mht@act.edu.in, [b] e-mail: nandhakumar.mht@act.edu.in

ABSTRACT

This paper surveys some of the principal research trends in social robotics and its application to human–robot interaction (HRI). Social (or sociable) robots are designed to interact with people in a natural, interpersonal manner – often to achieve social-emotional goals in diverse applications such as education, health, quality of life, entertainment, communication, and collaboration. The long-term goal of creating social robots that are competent and capable partners for people is quite a challenging task. They will need to be able to communicate naturally with people using both verbal and nonverbal signals. They will need to engage us not only on a cognitive level, but on an emotional level as well. They will need a wide range of social-cognitive skills and a theory of other minds to understand human behavior, and to be intuitively understood by people. A deep understanding of human intelligence and behavior across multiple dimensions (i. e., cognitive, affective, physical, social, etc.) is necessary in order to design robots that can successfully play a beneficial role in the daily lives of people. This requires a multidisciplinary approach where the design of social robot technologies and methodologies

INTRODUCTION

The way a person interacts with a social robot (or sociable robot) is quite different from interacting with the majority of autonomous mobile robots today. Modern autonomous robots are generally viewed as tools that human specialists use to perform hazardous tasks in remote environments (i. e., sweeping minefields, inspecting oil wells, mapping mines, etc.). In dramatic contrast, social (or sociable) robots are designed to engage people in an interpersonal manner, often as partners, in order to achieve social or emotional goals. The development of socially intelligent and socially skillful robots drives research to develop autonomous or semiautonomous robots that are natural and intuitive for the general public to interact with, communicate with, work with as partners, and teach new capabilities. Dautenhahn’s work is among the earliest in thinking about robots with interpersonal social intelligence where relationships between specific individuals are important. These early works pose the question What are the common social mechanisms of communication and understanding that can produce efficient, enjoyable natural, and meaningful interactions between humans and robots? Promisingly, there have been initial and ongoing strides in all of these areas. In addition, this domain motivates new questions for robotics researchers, such as how to design for a successful long-term relationship where the robot remains appealing and provides consistent benefit to people over weeks, months, and even years. The benefit that social robots provide people extends far beyond strict task performing utility to include educational, health and therapeutic, domestic, social and emotional goals (e.g., entertainment, companionship, communication, etc.), and more. We begin this chapter with a brief overview of a wide assortment of socially interactive robots that have been developed around the world. We follow with selected topics that highlight some of the representative research themes: multimodal communication, expressive emotion-based interaction, and social-cognitive skills.

SOCIAL ROBOT EMBODIMENT

Social robots are designed to interact with people in human centric terms and to operate in human environments alongside people. Many social robots are humanoid or animal-like in form, although this does not have to be the case. A unifying characteristic is that social robots engage people in an interpersonal manner, communicating and coordinating their behavior with humans through verbal, nonverbal, or affective modalities. As can be seen in the following examples, social robots exploitmany different modalities to communicate and express social-emotional behavior.

These include whole-body motion, proxemics (i. e., interpersonal distance), gestures, facial expressions, gaze behavior, head orientation, linguistic or emotive vocalization, touch based communication, and an assortment of display technologies. For social robots to close the communication loop and coordinate their behavior with people, they must also be able to perceive, interpret, and respond appropriately to verbal and nonverbal cues from humans. a number of socially interactive humanoid robots have been developed that can participate in whole body social interaction with people such as dancing, walking hand-in-hand, playing a musical duet, or transferring skills to unskilled persons. Their arms and hands are designed to exhibit human-like gestures such as pointing, shrugging shoulders, shaking hands, or giving a hug. Some of them are designed with mechanical faces to communicate with humans via facial expressions.

Whereas many of these humanoids have a mechanical appearance, android robots are designed to have a very human-like appearance with skin, teeth, hair, and clothes. A design challenge of android robots is to avoid the uncanny valley where the appearance and movement of the robot resemble more of an animate corpse than a living human. Designs that fall within the uncanny valley elicit a strong negative reaction from people. There are a number of more creature-like social robots that take their aesthetic and behavioral inspiration from animals. Given that people pet and stroke companion animals, touch-based communication has been explored in several of these more animal-inspired robots. Sony’s entertainment robot dog, AIBO [58.26], is a well-known commercial example. Other robots in this category have a more organic appearance, such as the therapeutic companion robot seal, Paro. Researchers have also chosen to design robots with a more fanciful appearance, melding anthropomorphic with animal-like qualities such as Leonardo, etc.. Many social robots are not overtly humanoid or zoomorphic, but still capture key social attributes, for instance, one of the best-known and pioneering social robots Kismet developed at the MIT Artificial Intelligence Lab. Kismet had a very expressive mechanical face with anthropomorphic features like large blue eyes.

Another example is the dancing robot Keep on developed by National Institute of Information and Communications Technology (NiCT) (Japan). This small yellow robot has a simplistic face and uses a classic animation technique called squash and stretch for expression of the body. Many mobile social robots have been fitted with faces to enhance social interaction. Some examples are the elder-care robot, Pearl, and the robotic receptionist Valerie with a graphical face on a liquid crystal display (LCD) screen, both developed at Carnegie Mellon University.

MULTIMODAL COMMUNICATION

Natural conversational ability is an important skill for social robots. Historically, even first-generation humanoid robots developed in the 1970s and 1980s (i. e., WABOT and WABOT-2) had primitive conversational skills modeled as simple combinations of speech input/output mappings. More recent examples are PaPeRo and receptionist robot (ASKA) that have conversational functions to work as information terminals. In natural human conversation, however, people send and receive nonverbal information that supplements linguistic information. These paralinguistic cues help smooth and regulate communication between individuals. The representative roles of paralinguistic information are as follows:

1. Regulators: expressions such as gestures, poses, and vocalizations that are used to regulate/control conversational turn-taking.
2. State displays: signs of internal state including affect, cognitive, or conversational states that improve interface transparency.
3. Illustrators: gestures that supplement information for the utterance. These include pointing gestures, iconic gestures,

ROBOTS THAT EXPRESS PARALINGUISTIC Information

In many cases, the same paralinguistic information can be conveyed through auditory or visual channels. However the characteristics of these channels have different properties. Sound has a strong property of attracting attention instantaneously. It can be used effectively for interruption, but it can also be disruptive and annoying. In addition, the timing of auditory paralinguistic cues is very strict. Overall, auditory paralinguistic signals are not suitable for continuous display. Visual paralinguistic cues are silent, the timing is a slightly less critical, and they can be used continuously. Hence, auditory and visual cues can be effectively used together to cooperatively convey the same information and emphasize it, or they can be used to convey different information simultaneously and contribute to efficiency. Thus, it is very important to choose the proper combination of modality and paralinguistic cues according to the situation. We provide examples below. Regulatory Cues Some of the earliest social robots displayed paralinguistic information to regulate interaction with people. Hadaly2 was the first robot to use mutual gaze as a paralinguistic cue to regulate conversation. When the robot and human achieved mutual gaze approximated using face recognition to determine when the human's face was facing the robot, Hadaly2 expressed readiness to commence conversation by blinking its eyes. Other examples are Kismet and Leonardo, which have implemented paralinguistic cues called envelope displays to regulate the exchange of speaking turns. Humans tend to make eye contact and raise their eyebrows when ready to relinquish their speaking turn, and tend to break gaze and blink when starting their speaking turn. These cues were shown to be effective in smoothing and synchronizing the exchange of speaking turns with human subjects, resulting in fewer interruptions and awkward long pauses between turns.

State-Display Cues

Other robots use state-display cues whereby the face or gaze of the robot is used to indicate its conversational or cognitive state. In general, this makes the robot's internal state more transparent to a person, so they can better predict and interpret the robot's conversational state and level of understanding. ROBITA used the tightness of its facial expression to indicate readiness to engage in conversation; a tight face was used to express conversational readiness, while a lose face communicated a lack of readiness to engage. Other state-display cues are back-channel responses. Human listeners use back-channel feedback (such as small head nods) to convey when speakers are successfully following the conversation. Response time for back-channel cues is very important because that cue is associated with the corresponding content. ROBISUKE employed finite-state transducer (FST) technology to achieve rapid understanding of the speech signal. This allowed the robot to resolve ambiguities in meaning and prepare its own responses even when the speaker was in mid-sentence. ROBITA provided humans with back-channel information via head nods and a tight face expression while listening. Another back-channel signal is an expression of confusion by the listener (verbal or nonverbal). This flags the speaker to stop and try to repair the broken communication. Robots such as Leonardo and ROBITA use facial displays of confusion when speech recognition fails in order to intuitively communicate to the human that he or she should repeat their last utterance. Similarly, gaze direction is a highly salient cue to convey the attentional focus of a robot. This is very useful if a human is trying to point out a particular object as a shared referent – such as pointing to an object before labeling it for the robot. These paralinguistic cues are useful for facilitating efficient conversational progress whereby errors and misunderstandings are identified and repaired immediately; for instance, Leonardo’s gaze behavior is used for performing envelope displays as well as for establishing shared attention with the human partner. Human subject studies have verified that Leonardo’s paralinguistic cues contribute positively to the transparency of the robot’s behavior and make the overall interaction more robust and efficient.

Illustrator Cues

A number of robots have implemented illustrator cues to direct the attention of a human. Often these robots use a variety of cues and the timing between them (such as gaze, head pose, pointing gestures, and conversational speech) to perform joint attention. In cases where a robot may be interacting with more than one person, the robot must properly take into account the location and orientation of the object, itself, and the individuals. ROBITA used such information to choose the proper gesture while considering the other people’s point of view.

Robots that Understand Paralinguistic Information

Humans readily express paralinguistic cues when interacting with robots just as they do with people. Consequently, conversational robots must be able to recognize and properly respond to these cues as well. This is a very difficult research challenge given the wide variety, subtlety, and timing of these human cues.

Regulatory Cues

A few robots can track a human’s paralinguistic cues to help regulate the conversation. The most common cues used for recognizing the end of the human’s speaking turn are mutual gaze (estimated using head pose to determine when human looks back to the robot) and paused speech. As a more sophisticated example, humans frequently provide short acknowledgement utterances (e.g., ‘uh-huh’, ‘um-hmm’, ‘huh’, etc.) as the robot explains something. These responses are either acknowledgments or acknowledgment-like repetitions, or ask-backs or ask-back-like repetitions. It is very difficult to distinguish these two kinds of utterances from the linguistic information as represented by the transcription of the utterance. The only way to distinguish them is by their prosody (not what, but how something is said). ROBISUKE could distinguish the utterance as either an acknowledgment or an ask-back from the prosody of utterance.

State-Display Cues

A number of robots are able to recognize and respond to state-display cues such as back-channel feedback nods, acknowledgement of an utterance, and attentional focus. One of the most robust systems for handling back channel feedback nods is Mel. A sophisticated head nod recognition system was developed whereby the robot could successfully distinguish small feedback nods from other kinds of head nods such as those that communicate agreement. Mel used this information to determine its own nodding behavior in order to be an appropriate response for the human. In a series of human subject studies, Sidner et al. found these paralinguistic cues to enhance the social engagement of the robot to people.With respect to recognizing successful or unsuccessful acknowledgement of an utterance, ROBISUKE used facial expression and prosody of the person’s utterance to make this determination. In a collaborative assembly task scenario, Sakita et al. presented a robotic system that used human gaze information to deduce the human’s intention of which object to operate upon next.

The robot used this information to choose an appropriate cooperative action such as either taking over for a human, settling a human's hesitation, or executing simultaneously with a human.

Illustrator Cues

Anumber of robots are able to recognize deictic gestures of a person conveyed either through pointing gestures or head pose. For example, Leonardo is able to infer the object referent in an interaction by considering a number of factors including pointing gesture, head pose, and speech. Brooks and Breazeal developed a deictic recognition system that enabled a robot to infer the correct object referent from correlated speech and deictic gesture. Interestingly, it was found that the accuracy of the human's pointing gesture is surprisingly poor. As a result, the deictic recognition system relies on coordinated speech and gesture information, with spatial knowledge provided by a three-dimensional (3-D) spatial database constructed by the robot using real-time vision, and a deictic spatial reasoning system. This sys- tem was successfully demonstrated on the dexterous humanoid Robonaut developed at National Aeronautics and Space and Administration (NASA) Johnson Space Center (JSC) where the human points to and labels a set of four bolts on a wheel to be fastened in order by the robot.

GROUP CONVERSATION

The ability to express and understand paralinguistic cues plays an important role in face-to-face conversation. The same is true for group conversation where a robot converses with two or more people. To continue a conversation, it is important for all conversational participants to understand who plays which role. ROBITA frames this problem with respect to information flow – to understand who is speaking to whom and when, and to determine each person's presence as a conversational partner. During group conversation, ROBITA tries to classify the participants as the speaker or as listeners. There is only one speaker at a time and the rest are listeners. The listeners are classified into a primal listener, to whom the utterance is directed, and secondly listeners, who observe the message exchange between the speaker and the primal listener. ROBITA discerns these roles by recognizing the face direction of the speaker. The person to whom the speaker is looking is recognized as the primal listener. To convey participation and improve its social presence in conversation, ROBITA tries to understand the conversational roles to look at the appropriate person. If the speaker faces ROBITA, then the robot recognizes itself as the primal listener and that the message is intended for itself. ROBITA faces the speaker when it is the primal listener. When ROBITA is a secondary listener, it looks at either the speaker or the primal listener.

COMMUNICATION IN COLLABORATION

Verbal and paralinguistic communication plays a very important role in coordinating joint action during collaborative tasks. Sharing information through communication acts is critical given that each teammate often has only partial knowledge relevant to solving the problem, each has different capabilities, and possibly diverging beliefs about the state of the task. For instance, all teammates need to establish and maintain a set of mutual beliefs regarding the current state of the task, the respective roles and capabilities of each member, and the responsibilities of each teammate. This is called common ground. Dialog certainly plays an important role in establishing common ground. Each conversant is committed to the shared goal of establishing and maintaining a state of mutual belief with the other. To succeed, the speaker composes a description that is adequate for the purpose of being understood by the listener, and the listener shares the goal of understanding the speaker. This communication acts serve to achieve robust team behavior despite adverse conditions, including breaks in communication and other difficulties in achieving the team goals. Humans also use nonverbal skills such as visual perspective taking and shared attention to establish common ground with others. They orient their own gaze and direct the gaze of their teammate through deictic cues such as pointing gestures in order to establish common ground. Given that visual perspective taking, shared attention, and the use of deictic cues to direct attention are core psychological processes that people use to coordinate joint action about objects and events in the world, robot teammates must be able to display and interpret these behaviors and cues when working with humans in a manner that adheres to human expectations. Breazeal et al. investigated the impact grounding using nonverbal social cues and behavior on task performance by a human–robot team.

In a human subject experiment, participants guided Leonardo to perform a physical task using speech and gesture. The robot communicates either implicitly through behavior (such as gaze and facial expressions) or explicitly through nonverbal social cues (i. e., explicit pointing gestures). The robot's explicit grounding acts include visually attending to the human's actions to acknowledge their contributions, issuing a short nod to acknowledge the success and completion of the task or subtask, visually attending to the person's attention directing cues such as to where the human looks or points, looking back to the human once the robot operates on an artifact to make sure its contribution is acknowledged, and pointing to artifacts in the workspace to direct the human's attention toward them.

CONCLUSION

In this chapter, we have presented some of the principal research trends in social robotics and human–robot interaction. We have relied heavily on examples from our own research to illustrate these trends, and have used excellent examples drawn from other research groups around the world. From this overview, we have shown that one of the most important goals social robots as applied HRI is the creation of robots that are human-compatible and human-centered in their design. Their differences from human abilities should complement and enhance our strengths. Their similarities to human abilities, such as computationally implementing human cognitive or affective models, may help us to understand ourselves better. We expect that in the coming decades many other researchers, especially young researchers, will actively contribute to the transition from today's robots into capable robot partners of tomorrow.

REFERENCES

[1] K. Dautenhahn: Getting to know each other – Artificial social intelligence for autonomous robots, Robot. Auton. Syst. **16**, 333–356 (1995)
[2] K. Dautenhahn: I could be you: The phenomenological dimension of social understanding, Cybern. Syst. **28**, 417–453 (1997)
[3] C. Breazeal: Designing Sociable Robots (MIT Press,Cambridge 2002)
[4] H. Miwa, A. Takanishi, H. Takanobu: Experimental study on robot personality for humanoid head robot, Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.(2001) pp. 1183–1188
[5] T. Tojo, Y. Matsusaka, T. Ishii, T. Kobayashi: A conversational robot utilizing facial and body expressions, Proc. IEEE SMC2000, Vol. 2 (2000) pp. 858–863
[6] J. Cassell, J. Sullivan, S. Prevost, E. Churchill: EmbodiedConversational Agents (MIT Press, Boston 2000)
[7] G. Hoffman, C. Breazeal: Robots that work in collaboration with people, Proc. AAAI Fall Symp. Intersec. Cognit. Sci. Robot.: From Interfaces Intell. Washington (AAAI, Menlo Park 2004)
[8] K. Dautenhahn, A.H. Bond, L. Canamero, B. Edmonds: Socially Intelligent Agents: Creating Relationships with Computers and Robots, Multiagent Syst.Artif. Soc. Simul. Organiz., Vol. 3 (Springer, Berlin, eidelber 2002)
[9] W.R. Picard: Affective Computing (MIT Press, Cambridge 1997)
[10] T. Fong, I. Nourbakshsh, K. Dautenhahn: A survey of social robots, Robot. Auton. Syst. **42**, 143–166 (2003)
[11] S. Schaal: Is imitation learning the route to humanoid robots?, Trends Cognit. Sci. **3**(6), 233–242 (1999)
[12] J. Solis, K. Chida, K. Suefuji, A. Takanishi: The development of the anthropomorphic flutist robot at aseda University, Int. J. Humano. Robots **3**(2), 1–25 (2006)

[13] Y. Ogura, H. Aikawa, K. Shimomura, H. Kondo,A. Morishima, H. Lim, A. Takanishi: Development of a new humanoid robot WABIAN-2, Proc. IEEE Int. Conf. Robot. Autom. (2006) pp. 76–81

[14] H. Miwa, K. Itoh, H. Takanobu, A. Takanishi: Mechanical design and motion control of emotion expression humanoid robot WE-4R, 15th CISM-IFToMM Symp.Robot Des.Dyn. Cont. ROM04–14 (2004)

[15] F. Tanaka, J.R. Movellan, B. Fortenberry, K. Aisaka: Daily HRI evaluation at a classroom environment: Reports from dance interaction experiments, Proc. 1st Annu. Conf. Human-Robot Interact. (HRI) (Salt Lake City 2006) pp. 3–9