IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received May 3, 2019, accepted June 4, 2019, date of publication June 10, 2019, date of current version June 26, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2921986

Proactive Robots With the Perception of Nonverbal Human Behavior: A Review

CHAPA SIRITHUNGE, (Student Member, IEEE),
A. G. BUDDHIKA P. JAYASEKARA, (Senior Member, IEEE),
AND D. P. CHANDIMA, (Senior Member, IEEE)
Intelligent Service Robotics Group, Department of Electrical Engineering, University of Moratuwa, Moratuwa 10400, Sri Lanka

Corresponding author: Chapa Sirithunge (ra-chapa@uom.lk)

This work was supported by the University of Moratuwa under Senate Research Grant SRC/LT/2018/20.

ABSTRACT Intelligent robot companions contribute significantly in improving living standards in the modern society. Therefore, human-like decision making skills are sought after during the design of such robots. On the one hand, such features enable the robot to be easily handled by its non-expert human user. On the other hand, the robot will have the capability of dealing with humans without causing any disturbance by the robot's behavior. Mimicing human emotional intelligence is one of the best and reasonable ways of laying the foundation for robotic emotional intelligence. As robots are widely deployed in social environments, perception of the situation or intentions of a user prior to an interaction is required to be proactive. Proactive robots are required to understand what is communicated by the human body language prior to approaching a human. Social constraints in an interaction could be demolished by this assessment in this regard. In this review, we incorporate various findings of human–robot interaction, social robotics and psychophysiology to assess intelligent systems which were capable of evaluating the emotional state of humans prior to an interaction. Second, we identify the cues and evaluation techniques that were utilized by such intelligent agents to simulate and evaluate the suitability of a proactive interaction. Available literature has been evaluated to distinguish limitations of existing methods and suggest possible improvements. These limitations, guiding principles to be adhered to and suggested improvements, are presented as an outcome of the review.

INDEX TERMS Interaction initiation, nonverbal cues, context-awareness, social robots, human-robot interaction.

I. INTRODUCTION

We live in a complex environment filled with various distinct phenomena. In the mean time, we revolve around other people we daily meet. We react to people and phenomena around ourselves in various means, based on the nature of a particular situation. Our reactions are visible to outside through behavior. Out of such reactions, words, expressions and physical movements are powerful drivers of human behavior [1].

Today, robots are deployed in our environments to make most of the tasks easier. Such robots are expected to have a general sense of the outcomes of their behavior in human-robot collaborative environments [2]. Many robots entering social environments are expert in only one or a few given specific tasks. Cleaning robots [3], [4], rescue robots [5], shopping assistants [6] and healthcare robots [7] are some examples for task-specialized robots which require lesser overall emotional intelligence. However with the deployment of robots in social environments, robotic systems have become a demonstrator of social and emotional interaction among humans and robots [8]–[10]. Therefore in the decades to come, the hospitality and emotional intelligence of artificial agents are expected to increase with their wide variety of social applications. Most significantly, the robots have to match between a particular situation and their emotional behavior [9].

Some situations urge a robot to take decisions regarding the pattern of interaction expected. Such situations request a conscious observation. Such robotic systems are already used in teaching, childcare and other social domains.

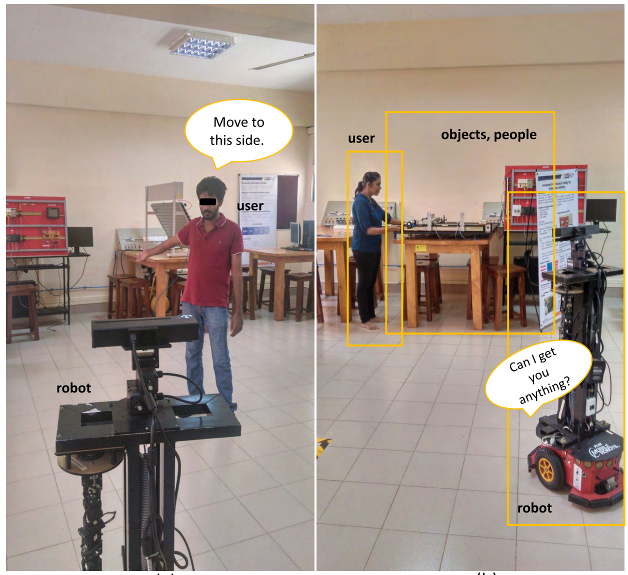
The associate editor coordinating the review of this manuscript and approving it for publication was Marina Gavrilova.

These systems still require development to adjust according to unexpected situations where human intentions change. Simulating and comprehending diverse social behaviors of humans is a challenging, and yet sought after feature to be engraved into a robot's personality [11]. In contrast, there are humanoid robots which can replicate typical human characteristics in the form of physical appearance, movements, words, facial expressions, etc. An example system in which robot's responses evolved by means of an adaptive behavior is presented in [12].

Human ascribes upon robots must be considered before planning this present robot's intelligence. For instance, creators in [13] researched how robots can enhance the personal satisfaction of senior citizens by being a partner to defeat loneliness. Here, individuals' mentality that robots are performing overwhelming undertakings but still lacking in intelligent conduct, could be changed with such an approach. For example [14], [15] are models for fall detection of elderly by robots. Thus, robots are acknowledged by numerous communities as a partner having the required capacities installed within. In such situations, following a decorum; essentially a "**robotiquette**" acknowledged by people is expected [16], [17]. Somehow numerous reactive techniques for human-robot domains have been set up so far, but yet not many proactive strategies were settled [18].

Preference-based collaboration with robots is a demanding viewpoint in human-robot shared situations [19]. In such circumstances, adopting supportive behaviors which rely on the insight, instead of playing out a requested task are required. Fitting such smart conduct is as critical as challenging. This is because of the multifaceted and complex nature of human conduct and perceiving such practices, outcoming the difficulties in innovation and environment. Accordingly human-level prediction of situations still needs enhancement [20].

A set of observable cues extracted from humans and their environment can be used as demonstrators for a perception model to simulate a human-robot scenario. Hence the level of interactivity within a situation can be determined. Nonverbal features of perceiving a situation are the most prominent and effective for an evaluation of the situation prior to an interaction. Understanding nonverbal behavior or the body language elevates decision making capabilities related to the interaction initiation by a robot, as the system outputs a measure of the emotional state in the human-robot encounter. Two such encounters are shown in Fig. 1. In the first instance in Fig. 1(a), the user gives instructions to follow. Hence the perception of the object placement, user's gestures and voice instructions will be adequate to perform the task. In contrast, Fig. 1(b) shows an occasion in which the user was unaware of the robot's presence. In both the situations, the state of the user is determined by the intentions of him/her as well as factors that exist in the surrounding environment. As Fig. 1 (b), objects and people make the environment, while intentions of the user are responsible for the factors within him/herself which might also affect the emotional state of that situation. In order to deliver a service in a polite manner, the robot



(a) (b)

FIGURE 1. **Example scenarios where the required level of situation-awareness differs. (a) shows a situation in which a human commands a robot to do a task (b) shows a situation where the user was engaged in a task. As the user was not attentive of the presence of the robot, robot tries to initiate an interaction. In such a scenario, facts from the surrounding as well as within the user may affect the emotional state of the interaction scenario [2].**

can adjust it's approach behavior, dialogs and other actions regarding interaction to match the situation. User's engagement in the task and interest for an interaction have to be predicted by the robot based on the factors visible in the scenario. Hence perception of emotional, social, psychological and other aspects of the situation become prominent during this type of scenario.

In this review, we aim at providing a detailed overview of the mechanisms used in modern artificial agents to gain emotional intelligence. We discuss the observable human cues used in support of such systems and evaluate to which extent such assessments by robots have been accepted by humans. We compare how humans perceive their surrounding when there are others around and the same concepts in human emotional intelligence in such scenarios have been adopted to evaluate the existing mechanisms used until now. Secondly, we compared earlier and modern approaches adopted by robots to evaluate observable, nonverbal human behavior. The review is concluded examining the most appropriate mechanisms to be adopted to gain emotional intelligence in human-robot collaborative environments. Furthermore, the areas where similar assessment of human behavior is required and challenges for such robots to overcome have been identified. Guiding principles and implications for future research are stated in the end. While this review analyses how robots perceive human behavior through observation, emotion representation and appearance of robots and the perception of its physical environment related to the encounter such as objects are beyond the scope of the review. On that account, this paper presents a review of literature on existing mechanisms and systems deployed to perceive nonverbal human behavior in human-robot encounters where a robot intends to initiate an interaction with its human user.

A. REVIEW PROTOCOL

The definitions and meanings of phrases used in the paper are listed below in order to provide a clear and concise understanding for the reader.

- **Situation-awareness:** Robot's intelligence upon the elements or events in an occasion encountered with a human. These elements are assessed analogous to time or space with the intent of finding meaningful relationships between occasions and behavior. The situation includes the humans and objects in the environment. The physiological, psychological, emotional, social and other aspects of the encounter are included in gaining awareness.
- **Nonverbal cues:** Features that could be observed from humans that do not involve direct verbal communication. These include bodily movements, facial expressions, details of appearance, postures, etc.
- **Interaction demanding or interaction readiness:** The extent to which a human prefers to have an interaction. The interest towards an interaction is measured by the interaction expected by a particular human. This interest is evaluated by using the observable cues displayed by that human.
- **State:** The current condition of an encounter. During HRI (Human-Robot Interaction), the parameters associated with the encounter between a human and a robot define the 'state' of that encounter. The state can be used to determine the level of interactivity required in an encounter.
- **Interaction initiation:** Start of an approach behavior or a conversation after observation of a particular situation. This is the phase during which the initiator gets the attention of the responder. In robot-initiated or proactive HRI, the robot is responsible for interaction initiation while the user is engaged.
- **Proactive robots:** Robots which identify the requirement of a certain situation and acts instantly without any instructions from outside.

The literature was selected from major indexed databases such as IEEE Xplore, Google Scholar, SCOPUS, ACM, Elsevier and ScienceDirect. Manuscripts published in peer-reviewed journals and conferences were considered to analyze during this survey. Unpublished or non-peer reviewed research articles such as technical reports, news, web articles, theses and dissertations were excluded from the study. In contrast, books and chapters which could support findings and definitions were included. Major focus was given to facts and findings established by journal articles upon conference papers. Only articles published in English were considered. Literature related to understanding and generating voice was excluded from this survey. In this review, proactive behaviors of industrial robots are excluded and only social robots are considered. Literature under Social Robotics, Human-Robot Interaction, Emotionally intelligent artificial agents and many more.

The paper is structured as follows. Nonverbal features of a human-robot encounter is discussed in section II. A brief discussion about features associated with human-robot encounters and what proactive robots should look for during HRI are discussed in section III. Section IV gives an overview about the current status of existing systems and directions of related research. Limitations identified in modern systems and possible improvements are discussed in section V. Finally the investigation is concluded in section VI.

II. NONVERBAL INTERACTION DURING HRI

Capacity of making inviting discussions at right occasions is indispensable in accomplishing a proactive behavior of a robot. Robots are required to have insight to decide when to cooperate and when not to. When starting communication, this natural or human-like conduct upgrades joint effort among robots and the non-master. So as to improve the union between human clients and robots, robots which can replicate human practices so as to assume the role of a close contact, for example, 'companion' and 'parental figure' are being produced [21]–[25]. However, the focus of these methodologies is restricted to the investigation and implementation of strategies that ought to be encapsulated into the robots. This helps robots to keep up a fluid interaction with their users. However, how does a robot identify key features to observe during a human-robot encounter?

To find answers to this question, researchers have been trying to establish common criteria to model human behavior. A large number of the current frameworks utilize facial expressions and body postures and additionally voice so as to decide the interaction demanding or the emotional state of a situation [26], [27]. In spite of the fact that voice can be utilized as a factor which demands interaction from an outsider, this is just conceivable after the user speaks. That is, after the user initiates the interaction. Meantime, a few methodologies have been introduced to assess the expressions and body postures during an interaction so as to judge user's likeliness towards an interaction with the robot [28], [29]. However, these realities can be assessed simply after the commencement of the interaction. Perceived information is used to decide upon the continuation of interaction flow thereafter. In any case, postures are useful in distinguishing the interaction demanding of the human subject before beginning a conversation.

Picard *et al.* suggested using haptic interaction tools to evaluate affective interactions [30]. This study uses body measures which are able to provide additional insight into an emotional state of a user without relying on an individual's cognitive assessment of the emotional state. But there is a rare chance for users to touch a robot (physical interaction) before an interaction and even they do, then it can be considered that the interaction was initiated by the user by means of touch, not the robot.

TABLE 1. Summary of the literature discussed in Section II.

Existing systems for human-robot nonverbal interaction	Key features	Highlights
[21]–[25]	The importance of human-like intelligence is identified. Application of robots with emotional and social intelligence is promoted.	Robot plays the role of a close companion rather than just being a 'machine'.
[26]–[29]	As most of the systems do, verbal responses, bodily postures and facial expressions which contribute to the emotional state of a situation were analyzed.	A limited number of observable cues from a situation were assessed by the robot.
[30]	Haptic perception of an encounter was introduced.	Emotional state of physical human-robot interaction (pHRI) was evaluated.

The above-discussed systems consider a limited set of cues from its surrounding to valuate the possibility of an interaction. Furthermore, integration of spatial and body factors are not adequately deployed in identifying the user behavior in these approaches. Restrictions for interaction must be parameterized and analyzed before the engagement with a user. For example, if the user was doing a certain task which involves rapid movements, it is reasonable for the robot to interpret the occasion as an 'engaged user'. On the other hand, if the user was standing, adopting very slow movements, that situation can be interpreted as a situation with a high interaction demanding where it is okay for the robot to interact with its user.

Which factors to be selected to analyze a situation is an area which requires attention. As a situation consists of the robot and its surrounding, factors from all these three aspects will need to be analyzed. Even so, development of a criteria which can bring all these factors to a common pool is still confusing and challenging. Difficulty of quantifying emotional factors for the analysis can be stated as the reason which retards the progress in respective field of research.

A summary of the discussed literature is given in TABLE 1.

A. WHAT TO OBSERVE?

Navarro and Karlins [1] reveal fascinating trends in nonverbal human behaviors, or the body language. In this way, transmission of information is achieved through body-based behaviors such as facial expressions, gestures, touching, physical movements, posture, and body adornment, etc. Nonverbal behavior comprises approximately 60-65% of all interpersonal communication or an interaction [31].

In order to derive intentions upon a certain situation, the robot has to be eclectic about the factors particular to that unique situation. A situation between a human and a robot consists of the *robot* itself, the *user* and the *environment* around the robot and human, as shown in Fig. 1. When the robot intends to perceive such a situation, it first has to identify interactive factors within itself and the user as well as which are in the environment. Factors within the *robot* itself include the dialog patterns the robot generates, maintaining an interactive distance in between, and displaying appropriate behavior, etc. Factors within the *user* will be numerous, but emotions, social norms, beliefs, personality traits, user's current activity and other psychophysiological factors contribute majorly to decide the level of interaction readiness within a human. Objects and other humans in the surrounding, obstacles, etc. make the list of factors in the *environment* which have to be perceived by the *robot*. Factors which are exhibited by the *robot* such as the personality traits, human-likeness, appearance, etc. were not considered within the scope of this review.

Psychological state is the reason behind human behavior. This psychological state is displayed to the outside through both verbal and nonverbal behavior. Body-based behavior is the result of cognitive processes developed in human brain. Behavior can be analyzed as an interplay of mental states and actions. Simply, thoughts and emotions provoke actions. In addition, cognitive elements such as facial expressions, verbal phrases, etc. fall under 'behavior' which includes both verbal and nonverbal aspects. Furthermore, brain activities such as internal states of mind, cognition and emotions are responsible for one's actions. Proper interaction between brain activity and actions, not only makes him perceive the world around him, but also enables the others around to perceive him. Fig. 2 shows how the behavior of the user changes in the presence of the robot.

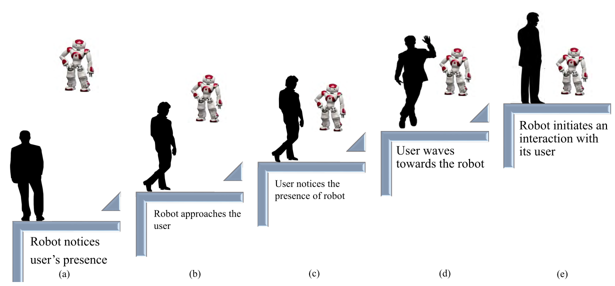


FIGURE 2. How the user behavior changes in the presence of robot as in [2]. (a) Initially only the robot is aware of the mutual existence of the user and robot (b) The user happens to see the robot as the robot approaches him (c) User's reaction for the approaching robot was looking at it (d) User waves hand towards robot (e) As the user responds to the robot in a friendly manner, robot decides to go closer and initiate a conversation.

Behavioral responses of humans can either be voluntary or sometimes involuntary. Furthermore, many involuntary behaviors are nonverbal. There is a number of psychological theories behind both voluntary and involuntary human behavior. Among these various theories, the theory of planned behavior and the theory of reasoned action map human actions and their thoughts in a rather reasonable and a justifiable basis [32]. Hence a method to perceive and combine these factors associated with a situation to predict the emotional state of a certain situation is required.

Introducing an established relationship between these various factors which are difficult to quantify, has been challenging at all the times. This is the reason to develop systems which evaluate only a limited number of factors so far. Often these systems either evaluate human factors or environmental factors alone.

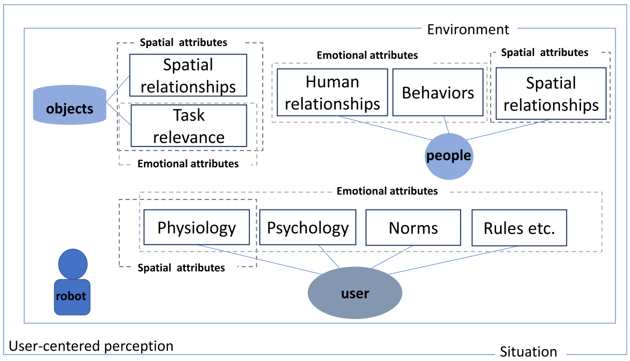


FIGURE 3. User-centered perception of a situation is shown. Each aspect; user, robot and environment is subdivided into smaller elements which affect the situation. For example, factors within the user are categorized into psychological attributes, physiological attributes, social norms and rules etc. Factors in the environment are divided into people and objects. Both spatial and emotional attributes of each aspect are considered.

B. COMMON ENCOUNTERS

The overall scenario addressed during the review is depicted in Fig. 3 and the situation was perceived by the robot in a user-centered manner. The user was not intended to adjust his/her behavior to fit the capabilities of the robot as in the past. Instead, the robot was allowed to adapt according to the perceived events around the user and other related subjects. Hence perception of factors external to robot is crucial in a user-centered design similar to the human-robot interaction scenario. Such situations are often encountered in social environments, domestic settings, museums, shopping malls, elderly and childcare etc.

1) USER-CENTERED DESIGN

User-centered design confronts designers to mold an interface around the capabilities and requirements of the operators. Rather than displaying information that is gathered around sensors and technologies which produce it, a user-centered design unites such various information in ways that set the goals, tasks, and requirements of the users. As a result of the user-centered design, we can extraordinarily diminish blunders and enhance productivity without requiring significantly new technical capabilities. Alongside user-focused structure, also comes enhanced user acceptance and satisfaction as a side advantage, by removing much of the frustration upon limitations of present technologies. User centered designs provide measures to support humans and then humans will work better with robots. The requirement of user's adaptation according to the limitations in technology is abolished during this approach.

In order to avoid people calling a robot explicitly, proactive robotic systems which can offer their help voluntarily are required. Therefore the user will not have to formulate his/her behavior to suit the capabilities of the robot, as the robot can read the intentions of its user. Simply, the proactivity of the robot reduces the user's effort.

A robot with cognitive skills engraved into its personality traits is more reliable [33]. Furthermore, humans display a more open and friendly behavior in front of a robot than another human. It is also upon robots that humans show more rejection when dissatisfied or disturbed by their behavior [2]. Hence situation-awareness of user behavior will mitigate incidents in which robots being rejected by their users.

In dynamic environments, decision making is largely reliant on robot's awareness- a constantly evolving picture of the state of the environment. This awareness drives decision making and consequently the performance of the robot [34].

2) REQUIREMENT OF PERCEPTION OF BEHAVIOR

The key points that have initiated the development of behavior-cautious robots and the requirement for such systems can be summarized as follows.

a. With the widened utilization of service robots in social environments over the past few decades, intelligence in such robots have to be developed to cater to human needs in close encounters. These systems should not make disturbances to humans through their behavior.

b. In most robotic systems, there are limitations in emotional intelligence than in its physical capabilities and efficiency. In contrast, emotional intelligence plays a major part in social environments. Monitoring human behavior is required before the robot initiates an interaction with a human, unless the individual requested for a specific service.

c. Users have to restrict their behavior when a robot lacks the capability to perceive their situation. For instance, if a robot invades the personal space of a user when approaching that individual, he/she has to limit his/her movements so as not to hit the robot. Such inaccuracies could be avoided if the robot perceives its surrounding in a human-friendly manner.

d. Situations which are favorable for a robot to initiate an interaction can be predicted through emotional intelligence of a robot. It will further enhance the relationship between humans and robots. In addition, robots will be accepted by its users for a longer duration as well.

e. Reliability of the robot by means of behavior can be improved by gaining emotional intelligence. Subsequently the events in which people getting disturbed by verbal or proxemic behavior of a robot, can be decreased. In addition, perception of human behavior has many applications other than Social Robotics. Caretaker robots, healthcare

TABLE 2. How cues displayed by humans vary in different scenarios.

Observable cues	Method of evaluation	Occasions
Facial Expressions	Existence of expressions	Domestic environments, Museums
Gestures	The types of gestures used	Rescue, Service applications
Voice responses	Meaning of the words	Social and academic environments
Pose and posture	Changes in pose and posture	Domestic and social environments, healthcare, Exhibits, Shopping malls
Bodily movements	Movements adopted at the particular instance	Domestic environments, Exhibits, Shopping malls
Gaze	Existence of gaze, aversion of gaze	All kinds of interactive scenarios

robots, rescue robots and robots deployed in extreme conditions such as disaster sites can make use of this capability to track and identify humans and their intentions to complete the robot's specified task accordingly. Furthermore, it's important to identify human behavior in order to generate most appropriate and timely responses during emergencies.

Interim Summary: Existing frameworks to interpret human behavior consider only a limited number of observable human cues. Therefore only a few aspects of a situation were effectively perceived by these mechanisms. On the other hand, establishing a common criteria to assess a situation, considering all the aspects: human, robot and environment, is still confusing and lacks conceptual basis. The key features of humans that could reveal their internal state of mind have been identified. In addition, different attributes: emotional and spatial, during a human-robot encounter have also been identified. This will create a user-centered perception of a robot, which will demolish the requirement of the human user to adopt a restricted behavior to accommodate the robot's limited perceptive skills. Therefore it has been identified that intelligent and adaptive decision-making skills in dynamic environments make robots more appealing and effectively performing in human-robot collaborative environments.

III. BUILDING ATTENTIVE ROBOTS

Theory of Mind has been developed in order put various attributes in mind together, to reason out a certain behavior and it has often been used to evaluate abnormal or awkward behaviors which have been deviated from the accepted level of behavior [35], [36]. The Theory of Mind has been used in several architectures to perceive human intention associated with a specific situation [37], but these approaches still estimate user's preferences based on a limited number of cues from its environment. Henceforth adequate evaluation of a situation is critical for emotional intelligence.

A. STUDY OF HUMAN BEHAVIOR

A human's intention alters involuntarily upon the factors that prevailed already in the surrounding. This perception will be based on various parameters including the individual's beliefs. Hence, the user's reaction upon an interaction which has been initiated by the robot will take different forms depending on the scenario. Responses that are most likely to be displayed from a human during a human-robot interaction scenario, can be used to assess human behavior during the study [38]. Such observable responses are listed below.

- Gaze - Maintaining or returning to original gaze e.g: looking at the robot and/or looking away
- Gestures - Using mainly hand gestures e.g: waving hand, calling in
- Postures - Changes in existing posture e.g: sitting to standing posture
- Utterances - Verbal responses e.g: "Hello", "May I get you something?"
- Movements - Random or intentional movements associated with activity
- Expressions - Facial expressions e.g: smile, disgust, frown

These responses devote to perceive and evaluate attitudes, attention, expectations, subjective norms and perceived behavioral control mainly as explained by the theory of planned behavior. Common encounters with the above behaviors are listed in TABLE 2.

The basic idea of this study is to evaluate the requirement of providing the robot with the ability of understanding situations or 'cognition'. Hence the cognition in situations will relate the connection between human's state and his/her behavior. This facilitates a dynamic interplay of flexibility and adaptation in robot.

B. ENVIRONMENTAL FACTORS

Orientation and locomotion are profound collaborators in 'interest' during an interaction among humans. This fact remains the same during human-robot interaction as well. Locomotion and attribution of body-based movement as an interpretation of own intentions are committing factors during HRI [39]. [40] provides an example of proactive obstacle avoidance in dynamic environments. Such systems are examples for intelligent agents with situation-awareness based on spatial behavior of subjects. Spatial arrangement of the two conversant is an important fact to determine the interactivity of a situation. Furthermore, the number of outsiders and placement of objects affect both directly and indirectly towards the emotional state of a situation. According to Fig. 3, the task of a user and approach behavior of an outsider are influenced by the placement of objects. Similarly, the relationships between people and behavior of the people in the surrounding further affect the responses of a human to a certain situation. Simply, all the subjects in the environment which are part of a person's cognitive mapping have an impact on his/her behavior [41]. Perception of natural environment such as symbol anchoring [42] and tracking dynamic obstacles [43] accounts for a proper understanding of the environment. However we exclude literature about robot's perception upon the *environment* and *robot* itself in this review.

C. FACTORS WITHIN THE ROBOT

Certain features and practices encapsulated in robots have an effect on individuals' willingness to participate in at least a brief interaction with the robot. Work explained in [44] has displayed a lot of social standards for robot conduct (a 'robotiquette') which is convenient and agreeable to people. As indicated by that, the conceptual space of HRI expects a robot friend in a home to 'do the correct things' and meet their expectations comfortably. Moreover, constant execution of the robot which pursues human social traditions and standards are bound to be acknowledged for a long span by its users [45].

In [46] the robot is exemplified with the ability to express emotions with a humanoid face and demonstrate attention by turning towards the subject. Creators have theorized that these highlights were non-negligible prerequisites for a viable social communication between a human and a robot. Anyway this framework has ignored enhancing the capacity of the robot to predict or realize user situation before using the previously mentioned two highlights while having an interaction. This approach gives an example of socially intelligent behavior of the robot as well. Even so it focuses on the behavior of the robot rather than that of the human. As per the model proposed in [47], **perception** and **evaluation** are constantly considered as critical attributes in human-robot interaction.

Authors in [48] present a framework for a mobile robot to initiate a dialog with a person and create an engagement. Such engagement was used to improve the performance of the robot's face recognition module with the help of the person. This robot further used body orientation to convey its interest for an interaction and, verbal responses and nonverbals such as gestures and arm and neck movements were used in support of the approach behavior. In this approach, the robot proactively seeks a human's help to learn its adjustments required for proper face recognition. A proactive behavior was adopted to improve the perception capabilities of the robot by enabling the human to teach it. Otherwise violation of user expectations when the robot asks help when the user is busy, will be unavoidable.

Interim Summary: Psychological theories such as the Theory of Mind and some previous studies in HRI identified various observable human cues related to movements and behavior that display human interest in a particular subject. Gaze, gestures, posture changes, involuntary movements, verbal responses and facial expressions to name a few. In this review we exclude voice responses, as a conversation opened by the user cannot be a part of robot-initiated interaction. In addition, there are environmental factors such as objects, distancing and location, and factors within the robot such as the personality and the physical appearance which will affect the emotional state of a human-robot encounter. As identified from this evaluation, a robot having a general sense of these attributes will receive a considerably higher attention as well as acceptance from its users.

IV. CURRENT STATUS: INTELLIGENT AGENTS WITH SOCIAL EMOTIONAL INTELLIGENCE

A. EARLIER APPROACHES

At the point when these robots are required to perform as domestic companions, starting a conversation at right events is very demanding among users. A great part of the users lean toward interaction by means of voice or friendly conversations [49]. Such insightful behavior upgrades the attachment and connection between the user and the robot [21]–[25]. In this way, gaining that emotional intelligence is a critical viewpoint concerning social situations.

There have been various psycho-physiological approaches to gain emotional intelligence through perception of behavior. Cognitive methods to understand the interactivity during a conversation have already been developed. These address the problem of identifying the intent of the conversational partner based on verbal cues and facial expressions. For example, in [50], the engagement of a human is assessed by the head-nod during a conversation.

In [51], a robot which was an insightful weight reduction mentor has been presented. This is an excellent example for behavior perception exemplified in the robot itself and subsequently has been deployed to reduce obesity among participants. In this case, the robot has been accepted for a long term interaction as well. Anyway this framework wasn't completely skilled to perceive rather general user situations other than monitoring physical health. By incorporating instinctive behavioral aspects in emotional situations into a robotic architecture, a higher emotional intelligence as well as a greater user acceptance can be ensured [52].

There are numerous automated frameworks to shape a discussion between a human and a robot, yet the ability of these frameworks are restricted simply after the initiation of an interaction, not before the interaction [54]. Mimicing human behavior is exceedingly applauded in cooperative discussions [55]. This methodology is prominent for keeping up interaction which has already been started, not prior to an interaction. Satake et. al in [53] enhances this scenario by taking various user cues into scrutiny. Forecasting of walking direction was valuable in estimating whether the user showed any interest for a conversation. 'User unaware' failure was abrogated in this method by moving toward the user before starting a conversation. In 'user unaware' failure, the user leaves the situation without realizing that the robot approaches to interact with him/her. Furthermore, in [53], a model to define an approach behavior for a robot to initiate a conversation with dynamic users has been proposed. The system was first intended to use in shopping malls to prevent 'user unaware' failure when a robot approaches a walking user. The approach model includes the following functions.

TABLE 3. Model of the approach behavior in [53].

Stage	Behavior of the robot	Failures abolished
Finding a target	Tracking a likely target	Unreachable/ rejected
Interaction at a public distance	Announcing the presence of robot and showing the intention to speak verbally	Unaware
Interaction at a social distance	Showing the intention of speaking nonverbally	Unsure

TABLE 4. A summary of the existing systems discussed in Section IV-A.

Earlier approaches to perceive nonverbal human behavior	Key features	Highlights
[49]	Interaction by means of friendly conversation.	Only voice responses were adjusted according to the situation.
[50]	Used facial expressions and verbal responses as well.	Robot's perception was not only limited to voice.
[51]	Behavior perception was used.	Monitored physical human behavior prior to an interaction.
[53]	Gaze and walking patterns were monitored to predict the intentions of a user before interaction.	Multiple cues were considered to evaluate a situation emotionally.
[53], [56]	Used gaze to identify a person's interest towards a certain encounter.	'Gaze' seems to play an important role in HRI.

1) Predicting the walking behavior of people
2) Choosing a target person
3) Planning its approaching path
4) Nonverbally indicating its intention to initiate a conversation

As a robot finds it difficult to recognize human gaze in a real environment, this method uses the body orientation of the subject and the robot as non-verbal cues which support an interaction. The robot's target was to approach humans and recommend shops in a selected mall. During this study, four occasions in which a robot may fail without or while interaction were listed.

1) unreachable- The robot did not get close to target person.
2) unaware- The person did not look at robot or did not listen to it.
3) unsure- The person recognized its presence and reacted but the robot did not respond correctly in time.
4) rejected- The person recognized its presence and it's greeting behavior, but did not start a conversation.

The model of approach behavior based on above is shown in Table 3. First the robot finds a target for interaction by predicting how people walk and estimating who can be approached. Then mutual distancing is determined. Here, human behavior was anticipated by means of walking direction and trajectories. Hence 'busy people', 'idlers', 'stationary people' and 'window shoppers' (just wandering around) were identified. Fig. 4 shows how these states were identified from the walking trajectory.

This mechanism was effective to predict the intention of dynamic users. Such approaches might not perform in smaller spaces such as domestic environments, as users adopt much less speeds and short trajectories in general.

Method proposed in [56] computed the 'degree of interest' of a user towards an interaction with a robot by evaluating

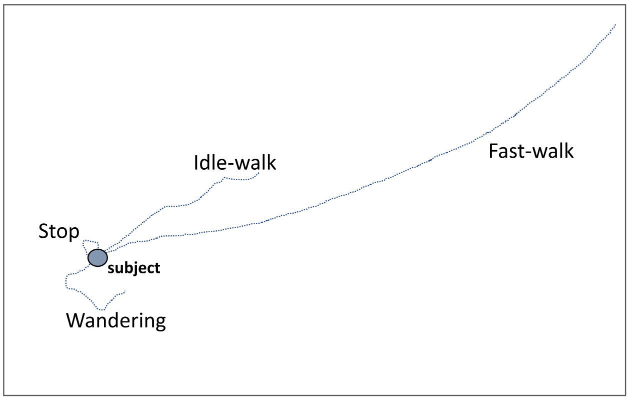


FIGURE 4. Classification of trajectories in different behaviors while walking as explained in [53].

the subject's attention and the distance from the robot to its subject by using a fuzzy logic based system. Head tilt angle was used as a variable which determines the attention of the subject. The output of the system was the degree of the user's interest. This system could perceive ambiguous situations with a higher accuracy with the help of linguistic variables that account for the fuzziness. However, head pose is not the only cue that demonstrates user's attention. Besides, a major concern while having a conversation is the distancing between the two conversant. Accordingly, the robot must decide upon the appropriate distancing for their communication before initiating a conversation. Personal zones suitable for such circumstances were presented in [57] and [58]. Hence a robot may choose which zone to enter while having an interaction, depending on the displayed interest of the user.

A summary of the discussed existing systems is given in TABLE 4.

Interim Summary: Most earlier approaches which intended to evaluate human attention or the interest towards an interaction focused on the face i.e. head nod or gaze. Verbal cues and facial expressions came next. In addition to movements, approach behaviors played an important role in finding human intention in dynamic environments such as museums and shopping malls. Some systems utilized the body orientation as a cue when recognizing gaze was difficult. Systems which used multiple cues as input, utilized fuzzy logic to evaluate vagueness in portraying intentions through behavior. Other approaches used predefined set of interpretations for most observable cues and appropriate robot's behavior. Previous studies further explored that higher emotional intelligence in robots increased user acceptance.

B. MODERN DEVELOPMENTS

A significant number of the past work have investigated the elements which influence human-robot interaction [44]. However exploring methods of initiation and maintaining smooth flow of a robot-initiated interaction, have been impeded due to practical and theoretical constraints. One under-investigated factor in this manner is perceiving nonverbal human behavior before communication. Once the robot is able to evaluate nonverbal human behavior at a certain scenario, it could evaluate the likelihood of initiating an interaction based on the encounter. In this manner, the impression of user situation is fundamental in such an event. Conversely, the absence of predictability and the transparency in numerous advanced mechanical frameworks have adversely influenced human's trust and dependence on robots [59]. A verbal approach for interaction is shown in [60]. During this approach, a robot which reacts dynamically to a visitor is deployed. Conversational opening is used as a critical influencer in maintaining visitor engagement. It further analyzed user's upper-body postures, facial expressions and head movements as demonstrators of engagement. However robot opened a conversation as a user approached. 'Pause' and 'restart' of a conversation by robot varied upon the variation of above demonstrators.

Law *et al.* [61] presented a similar mechanism to understand one aspect in this regard. A wizard-of-oz (WoZ) study has been conducted with the participation of humans to assess the curiosity level aroused in humans in the presence of assistive robots. The study confirmed that the human curiosity considerably diversified when the robot's intelligence is higher. The study further suggested that the social acceptance of a social robot increases when the robot can comprehend the very needs of its user and act according to the situation [62]. Results of the human study in [63] verified the fact that humans prefer user-adaptive dialogs while having a conversation even with a robot companion. Work proposed in [64] is an encouraging example for the growing harmony between humans and robots that are emotionally intelligent, as stated above. Users further expect the companionship grown upon interactive communication in between theirselves and robot, despite the service tasks most robots are designed for [65].

Likeliness of the robot being acknowledged as a conversational partner relies upon the surrounding and also the emotional state of the user. Moreover, it is imperative to investigate human inclinations towards a communication with a robot since human behavior before another human could take a different form in comparison to that before a robot. The human study performed in [2] was planned to investigate human responses towards an interaction with a robot in various circumstances. As indicated by this investigation, human reactions upon a robot-initiated interaction vary depending on their current activity. Consequently engraving social intelligence into a robot to perceive situations, is imperative before utilizing them in social environments.

In [66], nonverbal user engagement is approximated using initiator and responder gaze times, face orientation and feedback times for the two speakers as cues. Bodily postures of humans that are also an interesting parameter just as the gaze time were excluded in this method. A well known application for collaboration with the user could be found in a museum guide robot as explained in [67]. This research presented an effective mechanism for museum guide robots to reach visitors after an observation at a distance. Spatial formation of guests, body orientation and mutual gaze times of guests were considered as the critical cues during the observation by a guide robot. However there are circumstances in household situations in which spatial arrangement of the two parties and body orientation do not demonstrate any interest for interaction although the user intends to interact. Henceforth, the technique is not adoptable for domestic service robots.

Mimicking real-world HRI behaviors and simulating scenarios based on observable information are key features during human-robot collaboration (HRC) [68]. A predictive and adaptive system based on learning by demonstration is explained in [69]. This system is an example for robots with supportive behavior for the user but it does not identify many salient features which portray human intention. A situation-conscious model is proposed in [70] to improve the design and interactive capabilities of an industrial robot. These findings show that the impact of social and spatial environment have to be considered in order to design a context-aware robot.

In [71], an android system with the capability of monitoring noverbal behavior of humans. However this cannot generate physically appealing behaviors such as a robot. There has been systems which were capable to match the appearance and demeanor as well [72]. Even though such systems cover psychological and physical aspects in interaction, there should be emotional aspects as well in order to behave appropriately in social environments.

A context-sensitive approach to anticipate the human behavior while the human is followed by a robot has been proposed in [73]. Although this anticipated human motion while walking, prediction of human behavior during stationary situations becomes highly dependent of the emotional state and the task of that human. Hence anticipation of stationary situations is a complex process. There are CPU-intensive

TABLE 5. A summary of the modern systems discussed in Section IV-B.

Modern approaches to perceive nonverbal human behavior	Key features	Highlights
[2], [61], [63]	Surveys and WoZ studies contribute significantly to find human tendencies before developing such systems.	Teleoperation explores newer areas to focus on when developing a robot's intelligence.
[60]	Reacts dynamically to user's verbal responses.	Analyzed verbal responses, upperbody postures, head movements and facial expressions.
[66], [67]	Properties of gaze have been deployed to predict human intention.	Observations made at a distance were used.
[72]	Uses humanlike appearance to be used for a longer duration.	Humanlike attributes are a plus for robots as well.
[73], [74]	Human walk and activities were used to analyze a situation.	The fact that a human's internal state changes upon the task has been changed.
[77]–[79]	Parameters such as user's attention, engagement and level of interaction were determined in order to take decisions. Modern approaches try to use many cues from various aspects.	Much higher number of cues related to social, emotional, psychophysiological, spatial (and many more) have been used. e.g: gaze, gestures, postures, bodily movements, distancing, etc.

approaches to recognize human activities such as [74]. However less CPU-intensive techniques to monitor human behavior are admired when real-time decisions have to be taken. Requirement of lesser pre-processing in such techniques become advantageous in implementing them in real-time.

Perceiving emotional cues that shows affect is important in avoiding misbehaviors of robots and improving acceptance in human community [75]. Findings suggest that perception of nonverbal behavior positively impacts HRC and hence the understandability of the robot is increased as well [76]. A situated interaction method which uses behavioral cues such as proximity, velocity, sound and posture information is presented in [77]. A virtual companion adjusted himself according to the situation and outputs an engagement score at each occasion. Even so only the gaze behavior of the virtual companion was evolved based on behavioral cues of the user.

A summary of these modern systems is given in TABLE 5.

Interim Summary: Evaluating nonverbal behavior could determine the likelihood of initiating an interaction based on the features of that particular encounter. Conversation/ interaction opening is considered a critical influencer in maintaining user engagement and most of the modern approaches critically evaluate observable cues before making interaction decisions. Wizard-of-oz studies have been useful in exploring user tendencies in various encounters. Such studies have further discovered that users prefer adaptive behaviors adopted by robots, which enhance interactivity. Properties of gaze: gaze angle, gaze time, averted gaze etc. played an important role in determining nonverbal user engagement in modern approaches. Posture-based behaviors such as body orientation, posture changes, and body pose have also received attention in determining the state of a human in some literature. Some work considered walking patterns of humans to determine user-supportive approach behaviors before an interaction. In general, context-aware robots observe multiple behavioral cues and a measurable score, for instance 'engagement score' in [77], 'attention' in [78] and 'level of interaction' in [79], to determine the 'favorability for an interaction' of an encounter.

In comparison, modern approaches have focused on multimodal mechanisms which consider multiple cues from a human as much as possible, than earlier stages. Furthermore, evaluation of human cues in these mechanisms are moving towards learning algorithms (e.g: reinforcement learning, Bayesian networks, regressive models, artificial neural networks) other than logic (fuzzy logic, predefined set of situations and actions). In addition, newer approaches use cues from multiple aspects, for instance body orientation (spatial), gaze, gestures, pose changes, walking patterns (physiological) and facial expressions (psychological). The requirements and basis of both earlier and modern approaches of emotional intelligence have been the same over decades.

C. SYSTEMS WITH ADAPTIVE PERCEPTION OF SITUATION BASED ON HUMAN CUES

An intuitive, multimodal approach for interaction was introduced in a museum guide robot in [80]. This robot used audio and video information to shift attention among humans and to monitor its eye and hand gestures while speaking. This robot further interacts with multiple users at a time and uses a series of facial expressions based on users' interest. This method further determined focus of attention to be paid on each person based on the time when a person has last spoken, distance between the person and the robot, and its position relative to robot's front. Therefore this method can be stated as a first step towards perceiving a situation for adaptive social behavior.

Mead et. al proposed a method to evaluate the perception of distance in [81]. During this approach, robot used gestures and verbal responses of the user to determine the mutual distance. In [82] a probabilistic approach has been adopted to analyze the engagement in verbal cues and gestures used by a human during interaction. However both these approaches are possible only after initiating an interaction by the user. Hence the task performance of a human-robot team can considerably be improved by mutually understanding nonverbal cues responsible for the situation [83]. Therefore perception of nonverbal cues related to human behavior and relating various cues together to determine the state of the situation are the attention-seeking requirements during HRI.

An affective robot was deployed to interact naturally with customers in a shopping mall and to provide shopping information [64]. It further tried to build a rapport between customer and the robot by remembering customers who repeatedly visit. Findings were based on a wizard-of-oz experiment. This was an effort to identify which kind of robots were required by people to in their shopping routine. Features that have to be embodied in the robot by means of physical existence, interactivity and the capability of personalized communication. These three roles were defined on the following considerations.

1) Guiding- The robot was co-located with people so that people could ask questions such as where to shop.
2) Building rapport- As the robot becomes to be considered as a representative of the mall, a friendly behavior was expected from it. Therefore a personalized service was delivered to the customer.
3) Advertisements- The robot advertises for further shopping by attracting people.

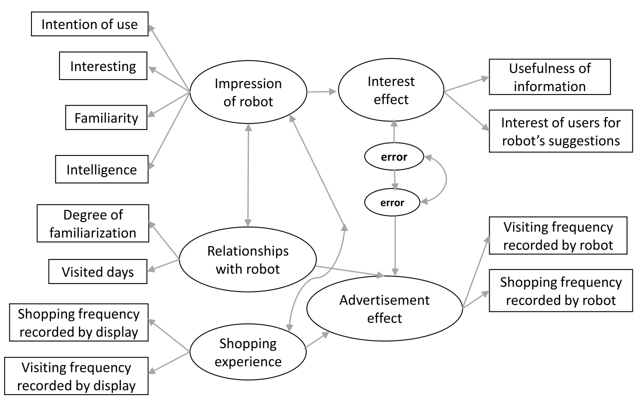


FIGURE 5. **Retrieved model about observed variables in [64].**

This decisioning process of the model is shown in Fig. 5. During this interaction process, people praised explicit identification of their names and other details by the robot. Furthermore they used a rather friendlier means of conversation after building a rapport. This is an example for a special scenario where a market-related situation perception was used. However actual human behavior will be more complex than in a shopping environment.

When robots are deployed in our ecosystems, they have to be self sufficient physically [84] as well as instincts in order to play the role of an active companion.

A framework to develop mixed-initiative approach for specifying the relationship between dialog structure and task structure using generic interaction patterns is proposed in [85]. Action-oriented dialogs were generated by a robot depending on the current task of its user. Objects from the environment and the task of a person were linked to find relationship in between, and hence conversation between the robot and the human was adapted accordingly during an object-grasping task. The work proposed in [86] is an example for a robot creating a map and maintaining a knowledge database through interaction with people in natural language.

In [87], a social robot was used to handle emotionally charged health care situations. According to the findings of this research, people's perception of robot was affected by how robots cope or how they think robots can cope with their emotions. In the end, improved intelligence of robots, like in these mechanisms, will increase people's acceptance of robots for longer interaction.

During the study of psychological benchmarks in HRI, in many forms of human-robot interaction there is almost nothing gained functionally by using a humanoid [88]. Furthermore, intelligence and capabilities become prominent in a robot upon its appearance. There are also contexts where the humanlike form might work against optimal human-robot interaction. For example, an older person might not want to be seen by a robot with a humanlike appearance when being helped to the washroom. In addition, people may dislike a robot that looks exactly human but lacks a humanlike behavioral repertoire [89]. Hence we do not consider the *robot aspect* during this survey of research on situation-awareness of robots.

Lack of intelligence pose issues for sociable service robots when user expectations are violated. This is inevitable when the robot is part of our physical environment and shares the world with us [38]. In a robot's view, human society will always be a challenging environment given its dynamic nature, richness in different scenarios, unpredictability, and uncertainty.

An effort to model interactivity of an encounter using four types of connection events which incorporate gesture and speech: directed gaze, mutual facial gaze, conversational adjacency pairs and back channels was taken in [90]. This uses initiator and responder gaze times to determine gaze and gaze is used with the other three parameters to determine user engagement during interaction. This is an example for using both verbal and nonverbal cues to perceive an interaction scenario. Adaptive speech control mechanism as a response for conditions in the environment has been proposed in [91]. However these are an example for evaluating a single or fewer number of cues from a situation.

User-awareness has been taken into consideration in [92] for safety and acceptance reasons. Spatial relationships, objects and dynamic gait behavior of a human were considered as features of the setting. Another mechanism for an approach behavior was proposed in [93]. But in this approach the tradeoff between user's and robot's gaze were used as

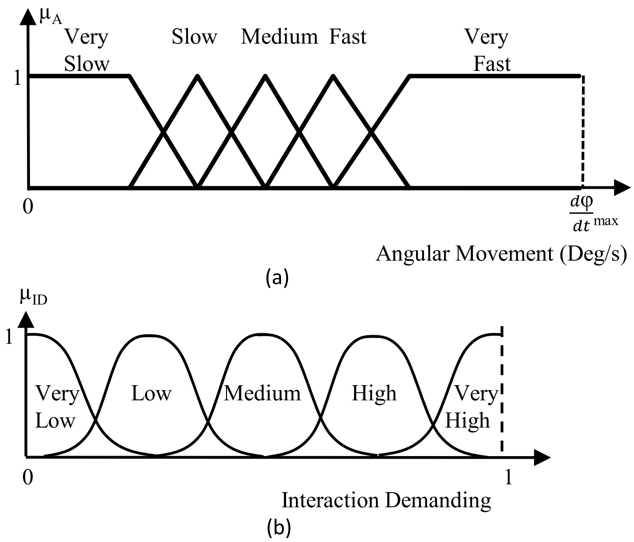


FIGURE 6. (a) and (b) represent the input membership function corresponding to angular movement and the output membership function which corresponds to the level of interaction demanding in [79].

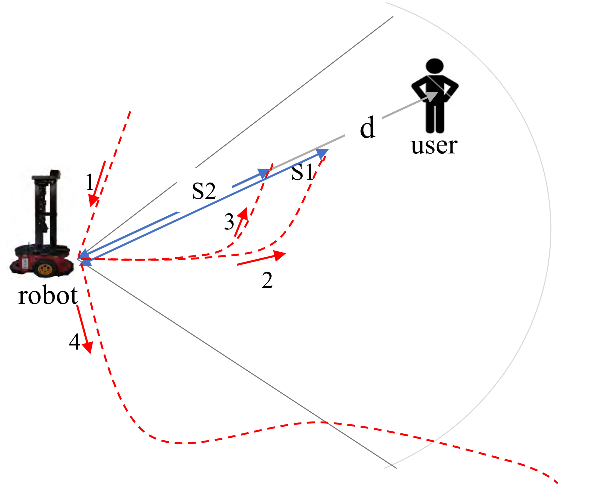


FIGURE 7. Routes taken by the robot during various encounters in [79]. Robot enters the scenario along route 1 and leaves along 4, if the attention of the user was 'low'. It approaches the user along route 2 and travels a distance of s2, if the attention was 'high' and approaches further along route 3 and travels a distance of s1 if the attention level was 'very high'.

a cue to evaluate the situation. Here, situation-awareness was used to determine approach behavior prior to an interaction. Another interactive conversational model which perceive space and verbal instructions has been proposed in [94]. A method to estimate the attention of a user is proposed in [79]. Authors have used the pose and speeds of specific angular joints of a human to predict the nonverbal interaction demanding of a human user. These two cues were evaluated with the help of a fuzzy logic based mechanism to derive a user's interest level towards a robot. Membership functions used for the evaluation are shown in Fig. 6. Decisions regarding interaction; whether or not to reach user and if reaches, the mutual distance to be kept in between were taken by this evaluation. Possible routes after decisioning in this approach are shown in Fig. 7. Even so the considered cues were not the only set of cues which associates nonverbal interaction demanding of a human.

Baraglia et. al in [95] presented how best humans collaborate with proactive robots. They have used human responses during a tray preparation task to evaluate this fact. This system used a Dynamic Bayesian Network (DBN) to anticipate environmental states in future and robot's actions to approach people. A specialty in this approach was that the robot switched between different assistive behaviors; helping when help is requested (human-initiated help), reactively helps when it realizes that help is needed. The robot evaluated a situation according to the defined criteria (robot-initiated reactive help) and proactively helped whenever it can help (robot-initiated proactive help). Their findings further elaborate that face gazes can be perceived as early cues for situation-awareness. To study these various help trigger mechanisms, an end-to-end system for executing joint tasks was developed. Every object had three predicates; 'object', 'position' and 'region'. Here, both the human and the robot were assumed to have one task-relevant action-pick and place. Here, 'pick' denotes the 'object' and 'place' denotes (x, y) coordinates of the goal. States of the objects were 'color', 'size' and 'location' as perceived by the robot. Actions were parametrized with the 'object' to be picked and the 'location' at which the object has to be placed. This model is given in Fig. 8.

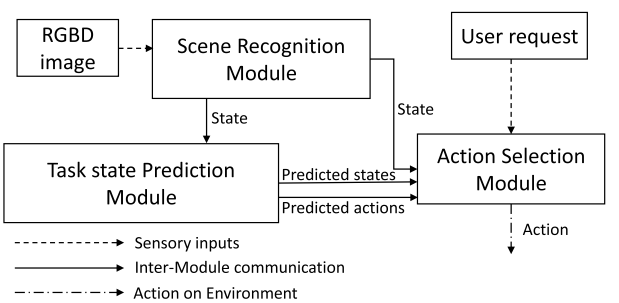


FIGURE 8. The model proposed for helping robots as in [95]: This model recognized the current environmental state and predicted possible future states with a dynamic Bayesian network. Then the actions to achieve the desired end-states were generated.

Actions were defined as a sequence of poses; pre-grasp, grasp and lift. Here, the actions of the robot were all related to manipulation of the object. Task execution was based on a two time-series DBN comprised of two multinomial nodes, environment states and actions. The policies involved

1) 'Human-initiated help' where the user explicitly says "Robot, can you help me?",
2) 'Robot-initiated reactive help' where the robot detects that predicted states were not reached, and
3) 'Robot-initiated proactive help' where robot takes human actions that might be in progress into account.

An interesting feature used in this approach is switching between these policies. However the system was capable of performing in pick and place tasks only, limiting its applications to only a fixed number of environments.

Museum guide robots observing visitors to find an appropriate situation to guide has been proposed in [67]. This perception was then used by the robot to locate himself in a spatial-oriental arrangement with a human. The purpose of this behavior was to allow humans perceive their participation. A transactional segment of space was used to monitor gaze and orientation of a visitor. A conversation was possible with a visitor only when both parties; robot and visitor establish a common belief to share a conversation. In this method, the robot believed that a visitor is interested in an exhibit if their face and body are oriented towards the exhibit for a certain duration. Similarly, their requirement to know further about the exhibit was determined by if they maintain face and body orientation towards the robot for a certain duration. A complex situation existed when there are more than one visitor in a single scenario. Therefore scenarios were categorized based on the directional gaze of each visitor. These were as follows.

1) When all the visitors are looking towards the robot- The robot turns its face and body towards them. The angle between the two vectors drawn from robot to each visitor; V_1 and V_2 marked in Fig. 9 (a) were calculated. Robot follows the median vector between V_1 and V_2 to approach visitors. It keeps a gap of 100-130 cm from the visitors. This complete process is shown in Fig. 9 (b), (c) and (d). After initiating the conversation, the robot moves to a position convenient to explain the exhibit.
2) If only some of the visitors are looking at robot- Robot pays attention to the ones who are looking at it.
3) When all the visitors are looking at the exhibit- Robot approaches them and waits for a favorable occasion to start conversation (robots wait for either scenario 1 or 2).

In this approach, the robot tried to maintain maximum level of interactivity without violating visitor expectations. But in social environments, human encounters will be more complex than those in a museum. Hence much more cues will be associated with user behavior.

A mobile robotic system which adaptively attend its user based on walking/sitting behavior of its user has been deployed in [96]. This system uses Finite State Machines (FSM) to model transitions of the user's state. The robot tracks a person's position and orientation to determine his/her state as illustrated in Fig. 10. The decision of the system based on these states was the path followed to approach the particular human. Here, three states: 'initial', 'walking' and 'sitting' were used by the FSM to model the transition of the person's state. Here, 'initial' refers to the person's state before recognition of the human. The adaptive attendance was determined as follows.

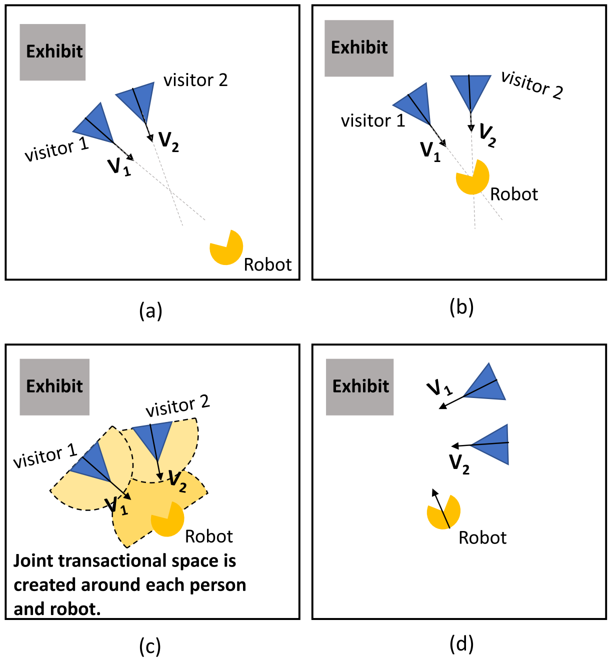


FIGURE 9. The occasion with two visitors looking towards the robot in [67].

FIGURE 10. The actual scenario and the estimated position and orientation of the target person as explained in [96].

1) Walking state: Robot follows the target with a gap of 1 m in between.
2) Sitting state: Robot moves to an appropriate waiting position by considering the structure of the environment and comfort of the person. This is further based on collision safety, comfortable attendance and social distance to the person. Hence a situation-aware attendance can be observed through this approach.

Accordingly, dynamic or static behaviors in humans have a great impact upon the interactivity of a situation. It considerably affect the approach behavior of the robot as well. Yuan *et al.* [97] evaluate the applicability of deep learning approaches to adapt to and predict comfortable proxemic behavior during interaction. The model estimated the discomfort when approaching its user. Distance between participant and robot, angle of approach, gender, age, previous experience with robots, preferred writing hand, and pet ownership was chosen as the inputs to the algorithm. Even though the research was about a comparison of different deep learning algorithms, the approach taken for the proxemic behavior was itself novel and situation-based. The model was capable of predicting proximity distances that were suitable in the context of HRI.

In [98], authors have proposed a system which used Regulatory Focus Theory, user's psychological state and game performance information to determine user's stress while playing a game. Hence the robot adapted its behavior accordingly. Hand movements of users were used to monitor stress level. Different kinds of body-based gestures and speech speeds of the robot were regulated based on the observations on the scenario. Furthermore this is an example for using body-based behaviors as a demonstrator of internal state of humans. [99] is and example for human's sensitivity towards empethetic and emotional features in a robot's speech. This fact can be associated with the overall performance of a robot as well in addition to speech alone.

In [39] the question 'How do the spatial distances and orientations of a user in relation to a robot vary throughout a cooperative task performed in a home-like environment?' has been addressed. Hence a spatial conduct for the robot was developed during this approach. This spatial management behavior was capable of active monitoring and dynamically reacting to each others' movement and position changes. This used the definitions of proxemics introduced by Hall [100] for interpersonal distancing and Kendon's F-formation arrangement for orientation between two persons [101]. The robot was intended to learn and find objects that were missing from its original location. For this purpose, the robot had to follow users who were willing to show the objects. Spatial formation of the human and the robot was analyzed during 'following the user', 'showing an object' and 'validating'. As different formations could be observed during each occasion, it could be deduced that there should be a perception upon each occasion during an interaction.

A computational model to recognize the engagement between human and robot is presented in [90]. This system used directed gaze, mutual gaze, conversational adjacency pairs and back channels as connection events in the decisioning process. Each connection event was analyzed on a timeline to find mutual occurrences. Information flow during recognition and decisioning are shown in Fig. 11.

In [102], body postures have been used as a mediator of affect during interaction. Here, an initial computational model to analyze postures and body motion to recognize

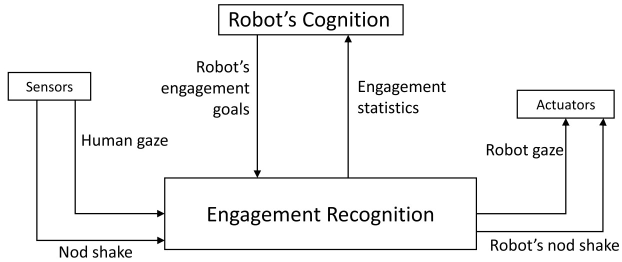


FIGURE 11. Retrieved model about observed variables as explained in [90].

engagement of children while playing chess with a robot has been introduced. Hence the robot act as a gaming companion but not a social companion. The engagement of the user was determined to sustain interaction throughout the game. A lateral view of the user was analyzed by a human to determine whether he/she was engaged in the game or not. Postural behavior was then analyzed using image processing techniques to relate the postural behavior with the engagement.

In [103], a robot that distributed flyers to pedestrians has been developed. This focuses on achieving the target by identifying the behavior of pedestrians and determining appropriate approaching mechanisms. This intended to approach a human being non-obstructive to the receiver. A video stream and positioning data of all the people in a selected area were used to calculate the detailed parameters. Following behavior types were used during implementation of the method.

1) Extend arm first and wait for pedestrian
2) Wait for the pedestrian and extend arm nearby
3) Extend arm first and approach pedestrian
4) Approach pedestrian and extend arm nearby

In addition, the timing of the processes such as approaching and handing over the flyer, and approach direction were also computed after a detailed observation on actual scenarios involving human-human interactions. Hence an entire distribution plan was deployed within a single situation.

A user's attention towards robot's presence was evaluated based on a number of parameters in [78]. This introduced a model to interpret the level of attention of the user as the robot approaches him/her. Upper body information extracted from RGB-depth data was used to approximate the level of attention of its user. This system deployed separate functional units assigned to extract information and analyze observed cues, estimate the attention level of the target subject and finally to take decisions regarding an interaction with that particular subject. Overall, it involved fuzzy evaluation of meaningful parameters which determine the attention of a human. This model used the spatial transformations in upper-body joints in vector format, which is important in quantifying slight movements which cannot be defined meaningfully in actual context. User's emotional state was determined by the variation of *gaze*, *gestures* and *changes in pose*. These major variables were subdivided into their properties and

TABLE 6. A summary of the adaptive systems discussed in Section IV-C.

Adaptive approaches to perceive nonverbal human behavior	Key features	Highlights
[80]–[82], [90], [67], [100]	Used gestures, gaze, spatial orientation of humans and verbal responses	Robot's tasks were adapted by these observations made at a distance.
[85], [86]	Dialogs and navigation were determined by updating a database.	Approach behaviors and conversations were the most adapted by a robot's careful observation.
[87]–[89]	Humanoid robots specialized in special tasks received more social acceptance.	We cannot deny the fact that the appearance of a robot has an effect on its acceptance among users.
[79], [93], [94], [96], [102]	User's gaze, gait, pose and movement speeds determined the state of the encounter.	Approach behavior was determined by analyzing the parameters associated with the encounter as the environment.
[95]	Verbal responses as well as task-oriented behavior were monitored.	Human activities during an object manipulation task were monitored to offer proactive help.
[97], [98], [103]	Dynamic and static behaviors, movement-based stress level were used as cues.	Emotional state associated with a situation was evaluated.
[78]	A large number of cues together with the features of these cues were utilized. e.g: Gaze (direction, duration, averted/not), Gestures (gesture speed, friendly gestures, duration of the gesture), pose/posture (changes in pose/posture)	Determined a user's attention towards a robot.

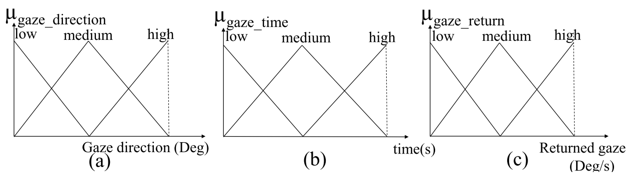


FIGURE 12. (a), (b) and (c) show the input membership functions of the gaze parameters: Gaze direction, gaze time and the level of returned gaze respectively as explained in [78]. These inputs determine 'gaze level' in Fig. 13 (a).

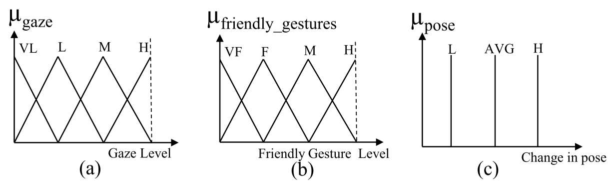


FIGURE 13. Output membership functions from *gaze* parameters, output from *friendly gestures* and *change in pose* in [78] are shown in (a), (b) and (c) respectively. These were then used as input parameters of a second fuzzy evaluation.

analyzed using fuzzy logic as shown in Fig. 12. *Gaze level* was determined by the 'gaze direction', 'gaze time' and 'gaze return' (whether the person looks away). Similarly, *gesture level* was determined by the 'gesture speed', 'gesture time' and 'happiness' where 'happiness' was defined by the usage of 'smile' and 'waving hand'. Finally, *attention* was estimated based on the variation of these variables; *gaze level*, *gesture level* and *changes in pose* as shown in Fig. 13. The overall assessment was based on two fuzzy systems and the output membership function used to determine the *attention level* is shown in Fig. 14. This system used many factors from the 'human'. Even so it has omitted factors in the environment which may have an impact upon the emotional state of the human-robot encounter. In addition, autoregressive models [104] can be used to model and predict behaviors

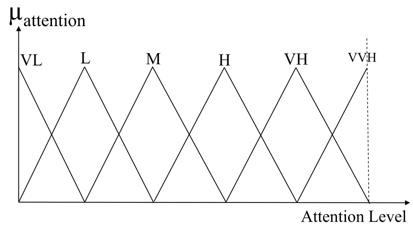


FIGURE 14. Output Membership Functions of second fuzzy evaluation in the system explained in [78].

with unexplained relationships. It further facilitates unsupervised simultaneous learning in human robot encounters [105].

Semantic maps also provide robots with an abstraction of the context. A clear understanding of objects and space relationships helpful in exploring emotional aspects associated with certain encounters. Hence the concept of semantic maps have a close association with human-robot collaborated environments. Therefore semantic maps will serve as a tool in combining the *human* and *robot* aspects with the *environment* in an encounter.

A summary of these modern systems is given in TABLE 6.

Interim Summary: Since a few decades, researchers were trying to use multimodal approaches to interpret the user behavior. Such approaches incorporate multiple cue from its subject of various aspects (psychophysiological, social, emotional, norms and rules, etc.). As most of the human behaviors are immeasurable, determining the internal state of mind based on such behaviors has been cumbersome. Therefore most existing work has deployed fuzzy logic, Bayesian networks, auto-regressive models, and machine learning techniques to evaluate nonverbal messages in our body language. Such approaches have been successful in interpreting emotions in behavior due to the difficulty of modeling vagueness and uncertainties associated with such inputs. Such evaluations were finally used to determine appropriate behaviors of the robot, such as approach behaviors, mutual distancing, type of conversation to initiate, emotions shown by the robot, etc. There have been methods incorporating multiple human cues which can reasonably represent the emotional state of an encounter. Still there are numerous cues remaining to be interpreted for their meaning and conveyed messages. Hence there is yet space for future researchers to improve the perception of intelligent agents to translate nonverbal human behaviors for affective and user-friendly HRI.

V. LIMITATIONS AND OUTLOOK

A. CHALLENGES IN IMPLEMENTING SITUATION-AWARENESS IN PROACTIVE ROBOTIC SYSTEMS

Following challenges have been disclosed during the study of present proactive robotic systems.

- Robust human detection and tracking systems to analyze gait and nonverbal behavior are not in par with advanced conceptual design of such systems.
- Establishing a relationship between intent and observable human cues has been challenging and lacks conceptual basis.
- Monitoring the environment in parallel with human behavior has been difficult and relating environment factors and human factors upon a certain behavior has been challenging.
- Deceiving and ironical human behaviors which do not match the shown intent with the real intent cannot be differentiated by existing robots.
- Bringing all observable cues to a combinable common platform in order to make interaction decisions, still needs further development. In other words, unwrapping the social and emotional attributes in physically observable cues still lacks conceptual basis.

Hence emotionally intelligent agents need development optimizing these challenges on the way.

B. IMPLICATIONS FOR AWARENESS THROUGH OBSERVATION

Much of the work suggest that there are factors in the environment and within the user itself, which affect user responses during a certain situation. Therefore the conceptual design of a robot's intelligence must consider these factors before implementing its task specific actions.

Most of the findings were based on a limited number of tasks selected from the environment and the user itself. In a real-life scenario, this number will be much higher than the number of factors considered in the present systems. Therefore a maximum number of parameters must be observed from the user and his/her environment before the decision-making process of a robot. Therefore these systems could not replicate all parts of the HHI (Human-Human Interaction) into the HRI scenario.

These methods were based on the assumption that people prefer the same rules of interaction with the robot as they do when interacting with humans. There can be certain cultures and social groups in which there are alterations regarding this fact. Hence such communities would react to robots in a different manner. In addition, behavior adaptation is as important as behavior monitoring in such a scenario. Several other factors which might influence interaction such as the gender, previous experience and familiarity with the robot were not considered within the context of many methods analyzed during this review. Furthermore there should be a common platform which can analyze psychophysiological, social, cultural, and other aspects of a situation. Finding relationships between these aspects and the emotional state of a situation is still found to be challenging. Unveiling the intentions behind the web of various human behaviors is another requirement in modeling cognitive models for social robots.

C. IMPLICATIONS FOR A SMOOTH INTERACTION

As users prefer their robots not to interrupt their usual behavior, the first design guideline suggested from these findings is to respect the preferences of humans by simply following their concerns. These 'concerns' can be determined by the factors considered in the study. This 'sense' of user situation further acts as an etiquette for the robot to fit well in social environments. This can be presented as the second design guideline for social robots.

The third design guideline is to extract information regarding the situation as much as possible. Considering a higher number of cues from the user and the environment increases the chance of an accurate perception of the situation. To perceive a number of such cues, the robot should acquire visual and auditory sensory information for an adequate duration. This will be the forth design guideline for a situation-aware robot. Establishment of a common criteria to evaluate different aspects of nonverbal cues is the fifth design guideline that can be deduced from this review. Such an approach could solve the problem of matching a psychophysiological behaviors and accurate interpretation of such behaviors. Ground for these implications are derived from the findings of this review.

D. POTENTIAL APPLICATIONS

Potential applications in which proactive robots with situation-awareness can be deployed can be listed as follows.

- As a companion in domestic service environments
- To assist people with special needs. In example such robots can support in children with autism, elders, sick and injured and robotic nurses which can tackle emergencies.
- As guides and to provide service. Such robots can be deployed in shopping malls, museums and cities.
- In rescue sites. These areas require management of the site and completing dangerous missions in areas unreachable by humans.
- To establish security by detecting theft and other harmful acts of humans which pose a threat to peace or the existence of general public.

Current status of methods used to gain situation perception in robots, their limitations and possible improvements are summarized taxonomically in Table 7. In Table 7, the key features of the scope covered by present systems and additional

TABLE 7. A summary of the existing systems and methods used for situation-awareness during HRI and suggested improvements in the future.

	Current Status	Possible Improvements
Scope	Observation and perception is limited to a few cues such as ● Spatial arrangements ● Components in the environment ● Human behavior oriented features ● Movements ● Facial expressions ● Postural information ● Voice commands	Time series assessment of social and psychophysiological cues in humans as well as factors in the environment. Maintaining a knowledge database is important to recognize and respond personally to each encounter. In addition, the robot has to maintain and constantly update an experience database. This will enhance robot's perception upon particular features in daily encounters.
Interaction Scenario	Many systems are based on verbal responses from humans in an encounter. Very few attempts have been taken to synchronize factors within and outside humans which affect equally on a situation. Rest of the work focus on evaluating only a limited number of cues that are easily modeled with each other to simulate state of a situation.	Extend the capabilities of present systems to perceive nonverbal cues as much as possible. Synchronize many cues which reflect different aspects of a situation to get the full picture. Adaptive synchronization of verbal and nonverbal cues to generate appropriate interaction mechanisms.
Adaptive Behavior	Information about following were taken into consideration for situation-awareness. ● Robot (appearance, mutual distancing, experience with the user) ● User (Body-based behavior, social norms, etiquettes) ● Environment (Spatial constraints, objects, people) Following cues were often utilized to perceive human-robot scenarios. ● Body-based movements [53], [73], [74], [78], [79] ● Posture [79] ● Facial expressions, Gaze, Response times [50], [53], [56], [66], [67], [79] ● Verbal responses [49], [50], [60] ● Spatial arrangement of conversant and objects [97], [98], [103] Following methods of evaluation were used. ● Dynamic Bayesian Networks [95] ● Deep neural nets [79], [96] ● Auto-regressive models [105] ● Fuzzy logic based systems [78] Performance evaluations of the systems were based on ● Comparison with human-human interaction ● User feedback scores [78], [79] ● Calculation of a determinant (eg: Satisfactory level of the user, engagement of the user) [78] ● Behavioral responses There is no common platform or a general theory which can explain the behavior of all the aspects (robot. user and environment) of a single encounter. Much of the experienced systems are based on a set of features associated with a single aspect e.g: facts regarding a particular user, environment, objects, robot's actions, user responses, but not all.	Synchronize multiple aspects with cues as much as possible. Find relationships between physical measurable quantities and emotional reactions of each. ● eg: Relating a gaze associated with a smile to the interest level of a human. Recognition of a task and the engagement by means of physical demonstrators such as body-based movements, facial expressions and objects associated with the task. Consider environmental and social factors as much as possible. Eg: ● Arrangement of objects and people ● Previous experience with people and objects ● Norms and etiquettes followed by the considered community Extraction of only the required cues, e.g: Exclusion of random behvaious like hand and leg movements. Adaptation based on previous experience Types of interaction used within a similar domain or the same user. Evaluate the possibility of using reinforcement learning to deveop a reward-based system to predict possible outcomes of interaction prior to an interaction. Evaluate the possibility of autoregressive models to synchronize cues of different aspects. Introduce an evaluation matrix and the level of interactivity in a specific situation. An extended database system has to be maintained to match all the complexities associated with a an encounter. Further development of cognitive theories which can explain complex encounters in general. Hence perceiving the setting and generating appropriate behaviors will be reasonably achieved by a robot.

features that could enhance the skills of such robots are given under 'Current Status' and 'Possible Improvements'. Next, the situations encountered during a typical HRI scenario at present are given under 'Interaction Scenario'. In that section, the requirement of situation-driven behavior is identified and approaches towards that behavior are pinpointed under 'Adaptive Behavior'.

VI. CONCLUSION

Enter of artificial agents in to human environments explored new horizons in human-robot interaction. In the present review, we provided an overview of the robotic systems with the perception of human behavior. Such perceptive intelligence was used to generate situation-cautious responses rather than delivering only what is requested by a human. Many mechanisms were deployed by proactive robots to evaluate nonverbal human behavior or the body language before the initiation of an interaction during a social encounter. Such evaluations were used to generate situation-friendly responses from a robot without disturbing its human users or violating their expectations.

Perception of nonverbal cues displayed by humans involved in a scenario was the prominent and most effective means of perceiving body language. Verbal cues are rather direct in delivering information to outside, but this is possible only after initiating an interaction by the human. On the other hand, behaviors such as facial expressions and smile which directly targets someone, can be deceiving and the person's real intentions may differ from what is shown to outside. Therefore nonverbal cues are more appropriate and less disturbing to a user in social encounters. Without going for direct interaction, situation-cautious behavior further gave a robot more human-like personality based on social etiquettes. Hence the ability to interpret the emotional state of a scenario and evaluating that scenario for the suitability of an interaction makes HRI more adaptive and meaningful.

Evaluation of a number of nonverbal observable cues is required for proactive interaction. Therefore existing such systems have been critically investigated in this review. Neuro-fuzzy based approaches, machine learning techniques and auto-regressive models have been most prominent in situation-cautious proactive robotic systems. The ability of such methods to model inexplainable relationship between human psychology and behavior has been the reason for wide usage of such approaches. Difficulty of integrating emotional states into a consistent user model mathematically or scientifically, and explaining reasonably the intentions of explicit human behavior, have retarded advancement of situation-aware robots.

Limitations of existing systems have been identified and possible future improvements have been suggested. This review provides principles and approaches for further development of perceptive intelligence. In summary, capability of existing robotic systems to make proactive decisions based on human behavior is far below that of a human. Furthermore literature related to robot-initiated interactive systems with situation-awareness is relatively scarce even though there is a great potential of development in this aspect.

REFERENCES

[1] J. Navarro and M. Karlins, *What Every Body is Saying*. New York, NY, USA: William Morrow Paperbacks, 2008.

[2] H. P. C. Sirithunge, M. A. J. V. Muthugala, A. G. B. P. Jayasekara, and D. P. Chandima, "A wizard of oz study of human interest towards robot initiated human-robot interaction," in *Proc. 27th IEEE Int. Symp. Robot Hum. Interact. Commun. (RO-MAN)*, Aug. 2018, pp. 515–521.

[3] J. L. Jones, N. E. Mack, D. M. Nugent, and P. E. Sandin, "Autonomous floor-cleaning robot," U.S. Patent 9 622 635, Apr. 18, 2017.

[4] A. K. Bordoloi, M. F. Islam, J. Zaman, N. Phukan, and N. M. Kakoty, "A floor cleaning robot for domestic environments," in *Proc. Adv. Robot.*, 2017, p. 19.

[5] L. Pineda, T. Takahashi, H.-T. Jung, S. Zilberstein, and R. Grupen, "Continual planning for search and rescue robots," in *Proc. IEEE-RAS 15th Int. Conf. Humanoid Robots (Humanoids)*, Nov. 2015, pp. 243–248.

[6] N. Doering, S. Poeschl, H.-M. Gross, A. Bley, C. Martin, and H.-J. Boehme, "User-centered design and evaluation of a mobile shopping robot," *Int. J. Social Robot.*, vol. 7, no. 2, pp. 203–225, 2015.

[7] T. Tanioka, K. Osaka, R. Locsin, Y. Yasuhara, and H. Ito, "Recommended design and direction of development for humanoid nursing robots perspective from nursing researchers," *Intell. Control Automat.*, vol. 8, no. 2, p. 96, 2017.

[8] J. Albo-Canals, A. B. Martelo, E. Relkin, D. Hannon, M. Heerink, M. Heinemann, K. Leidl, and M. U. Bers, "A pilot study of the KIBO robot in children with severe ASD," *Int. J. Social Robot.*, vol. 10, no. 3, pp. 371–383, 2018.

[9] J. Goetz, S. Kiesler, and A. Powers, "Matching robot appearance and behavior to tasks to improve human-robot cooperation," in *Proc. 12th IEEE Int. Workshop Robot Hum. Interact. Commun.*, Nov. 2003, pp. 55–60.

[10] S. Andrist, X. Z. Tan, M. Gleicher, and B. Mutlu, "Conversational gaze aversion for humanlike robots," in *Proc. ACM/IEEE Int. Conf. Hum.-Robots Interact.*, 2014, pp. 25–32.

[11] I. Leite, C. Martinho, and A. Paiva, "Social robots for long-term interaction: A survey," *Int. J. Social Robot.*, vol. 5, no. 2, pp. 291–308, 2013.

[12] J. Bongard, "Using robots to investigate the evolution of adaptive behavior," *Current Opinion Behav. Sci.*, vol. 6, pp. 168–173, Dec. 2015.

[13] H. A. Karim, A. M. Lokman, and F. Redzuan, "Older adults perspective and emotional respond on robot interaction," in *Proc. 4th Int. Conf. User Sci. Eng. (i-USEr)*, Aug. 2016, pp. 95–99.

[14] J. Y. Hwang, J. M. Kang, Y. W. Jang, and H. C. Kim, "Development of novel algorithm and real-time monitoring ambulatory system using Bluetooth module for fall detection in the elderly," in *Proc. 26th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (IEMBS)*, vol. 1, Sep. 2004, pp. 2204–2207.

[15] D. Stavrou, D. G. Eliades, C. G. Panayiotou, and M. M. Polycarpou, "Fault detection for service mobile robots using model-based method," *Auto. Robots*, vol. 40, no. 2, pp. 383–394, 2016.

[16] P. Liu, D. F. Glas, T. Kanda, H. Ishiguro, and N. Hagita, "A model for generating socially-appropriate deictic behaviors towards people," *Int. J. Social Robot.*, vol. 9, no. 1, pp. 33–49, 2017.

[17] I. Pedersen, S. Reid, and K. Aspevig, "Developing social robots for aging populations: A literature review of recent academic sources," *Sociol. Compass*, vol. 12, no. 6, 2018, Art. no. e12585.

[18] M. G. Rashed, D. Das, Y. Kobayashi, and Y. Kuno, "A study on proactive methods for initiating interaction with human by social robots," *Asian J. Converg. Technol.*, vol. 4, no. 11, pp. 1–10, 2018.

[19] C.-A. Smarr, T. L. Mitzner, J. M. Beer, A. Prakash, T. L. Chen, C. C. Kemp, and W. A. Rogers, "Domestic robots for older adults: Attitudes, preferences, and potential," *Int. J. Social Robot.*, vol. 6, no. 2, pp. 229–247, 2014.

[20] J. D. Mayer, P. Salovey, and D. R. Caruso, "TARGET ARTICLES: 'Emotional intelligence: Theory, findings, and implications,"' *Psychol. Inquiry*, vol. 15, no. 3, pp. 197–215, 2004.

[21] S. Strohkorb and B. Scassellati, "Promoting collaboration with social robots," in *Proc. 11th ACM/IEEE Int. Conf. Hum. Robot Interact.*, Mar. 2016, pp. 639–640.

[22] G. Hoffman and C. Breazeal, "Robotic partners' bodies and minds: An embodied approach to fluid human-robot collaboration," *Cogn. Robot.*, 2006.

[23] M. Heerink, B. Vanderborght, J. Broekens, and J. Albó-Canals, "New friends: Social robots in therapy and education," *Int. J. Social Robot.*, vol. 8, no. 4, pp. 443–444, Aug. 2016.

[24] B. Robins, K. Dautenhahn, R. Te Boekhorst, and A. Billard, "Robotic assistants in therapy and education of children with autism: Can a small humanoid robot help encourage social interaction skills?" *Universal Access Inf. Soc.*, vol. 4, no. 2, pp. 105–120, Dec. 2005.

[25] M. N. Nicolescu and M. J. Mataric, "Learning and interacting in human-robot domains," *IEEE Trans. Syst., Man, Cybern. A, Syst. Humans*, vol. 31, no. 5, pp. 419–430, Sep. 2001.

[26] H. Gunes and M. Piccardi, "Bi-modal emotion recognition from expressive face and body gestures," *J. Netw. Comput. Appl.*, vol. 30, no. 4, pp. 1334–1345, Nov. 2007.

[27] M. Coulson, "Attributing emotion to static body postures: Recognition accuracy, confusions, and viewpoint dependence," *J. Nonverbal Behav.*, vol. 28, no. 2, pp. 117–139, 2004.

[28] M. P. Michalowski, S. Sabanovic, and R. Simmons, "A spatial model of engagement for a social robot," in *Proc. 9th IEEE Int. Workshop Adv. Motion Control*, Mar. 2006, pp. 762–767.

[29] C. H. Chen and P. S.-P. Wang, Eds., *Handbook of Pattern Recognition and Computer Vision*, 3rd ed. River Edge, NJ, USA: World Scientific, 2005.

[30] R. W. Picard and S. B. Daily, "Evaluating affective interactions: Alternatives to asking what users feel," in *Proc. CHI Workshop Eval. Affect. Interfaces, Innov. Approaches*, 2005, pp. 2119–2122.

[31] J. K. Burgoon and D. B. Buller, "Interpersonal deception: III. Effects of deceit on perceived communication and nonverbal behavior dynamics," *J. Nonverbal Behav.*, vol. 18, no. 2, pp. 155–184, 1994.

[32] T. J. Madden, P. S. Ellen, and I. Ajzen, "A comparison of the theory of planned behavior and the theory of reasoned action," *Pers. Social Psychol. Bull.*, vol. 18, no. 1, pp. 3–9, 1992.

[33] S. Petisca, J. Dias, P. Alves-Oliveira, and A. Paiva, "Emotional sharing behavior for a social robot in a competitive setting," in *Proc. 25th IEEE Int. Symp. Robot Hum. Interact. Commun. (RO-MAN)*, Aug. 2016, pp. 730–735.

[34] M. R. Endsley, *Designing for Situation Awareness: An Approach to User-Centered Design*. Boca Raton, FL, USA: CRC Press, 2016.

[35] B. Scassellati, "Theory of mind for a humanoid robot," *Auto. Robots*, vol. 12, no. 1, pp. 13–24, 2002.

[36] S. K. Card, *The Psychology of Human-Computer Interaction*. Boca Raton, FL, USA: CRC Press, 2017.

[37] O. C. Görür, B. S. Rosman, G. Hoffman, and S. Albayrak, "Toward integrating theory of mind into adaptive decision-making of social robots to understand human intention," presented at the Int. Conf. Workshop Role Intentions Hum.-Robot Interact., Vienna, Austria, Mar. 2017.

[38] C. Breazeal, "Social interactions in HRI: The robot view," *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, vol. 34, no. 2, pp. 181–186, May 2004.

[39] H. Hüttenrauch, K. S. Eklundh, A. Green, and E. A. Topp, "Investigating spatial relationships in human-robot interaction," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Oct. 2006, pp. 5052–5059.

[40] M. Göller, F. Steinhardt, T. Kerscher, J. M. Zöllner, and R. Dillmann, "Proactive avoidance of moving obstacles for a service robot utilizing a behavior-based control," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Oct. 2010, pp. 5984–5989.

[41] R. M. Downs and D. Stea, *Image and Environment: Cognitive Mapping and Spatial Behavior*. Piscataway, NJ, USA: Transaction Publishers, 1974.

[42] M. Weser and J. Zhang, "Proactive multimodal perception for feature based anchoring of complex objects," in *Proc. IEEE Int. Conf. Robot. Biomimetics (ROBIO)*, Dec. 2007, pp. 1069–1074.

[43] A. Sabnis, G. Arunkumar, V. Dwaracherla, and L. Vachhani, "Probabilistic approach for visual homing of a mobile robot in the presence of dynamic obstacles," *IEEE Trans. Ind. Electron.*, vol. 63, no. 9, pp. 5523–5533, Sep. 2016.

[44] K. Dautenhahn, "Socially intelligent robots: Dimensions of human–robot interaction," *Philos. Trans. Roy. Soc. B, Biol. Sci.*, vol. 362, no. 1480, pp. 679–704, 2007.

[45] T. Fong, I. Nourbakhsh, and K. Dautenhahn, "A survey of socially interactive robots," *Robot. Auto. Syst.*, vol. 42, nos. 3–4, pp. 143–166, 2003.

[46] A. Bruce, I. Nourbakhsh, and R. Simmons, "The role of expressiveness and attention in human-robot interaction," in *Proc. IEEE Int. Conf. Robot. Autom.*, vol. 4, May 2002, pp. 4138–4142.

[47] J. Scholtz, "Theory and evaluation of human robot interactions," in *Proc. 36th Annu. Hawaii Int. Conf. Syst. Sci.*, Jan. 2003, pp. 1–10.

[48] A. Garrell, M. Villamizar, F. Moreno-Noguer, and A. Sanfeliu, "Proactive behavior of an autonomous mobile robot for human-assisted learning," in *Proc. IEEE RO-MAN*, Aug. 2013, pp. 107–113.

[49] C. Ray, F. Mondada, and R. Siegwart, "What do people expect from robots?" in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Sep. 2008, pp. 3816–3821.

[50] C. L. Sidner, C. Lee, L.-P. Morency, and C. Forlines, "The effect of head-nod recognition in human-robot conversation," in *Proc. 1st ACM SIGCHI/SIGART Conf. Hum.-Robot Interact.*, 2006, pp. 290–296.

[51] C. D. Kidd and C. Breazeal, "Robots at home: Understanding long-term human-robot interaction," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Sep. 2008, pp. 3230–3235.

[52] R. C. Arkin, M. Fujita, T. Takagi, and R. Hasegawa, "An ethological and emotional basis for human–robot interaction," *Robot. Auto. Syst.*, vol. 42, nos. 3–4, pp. 191–201, 2003.

[53] S. Satake, T. Kanda, D. F. Glas, M. Imai, H. Ishiguro, and N. Hagita, "How to approach humans?: Strategies for social robots to initiate interaction," in *Proc. 4th ACM/IEEE Int. Conf. Hum. Robot Interact.*, Mar. 2009, pp. 109–116.

[54] B. Mutlu, T. Shiwa, T. Kanda, H. Ishiguro, and N. Hagita, "Footing in human-robot conversations: How robots might shape participant roles using gaze cues," in *Proc. 4th ACM/IEEE Int. Conf. Hum. Robot Interact.*, Mar. 2009, pp. 61–68.

[55] C. L. Sidner, C. D. Kidd, C. Lee, and N. Lesh, "Where to look: A study of human-robot engagement," in *Proc. 9th Int. Conf. Intell. User Interfaces*, 2004, pp. 78–84.

[56] E. Aguirre, M. Garcia-Silvente, A. González, R. Paúl, and R. Munyoz, "A fuzzy system for detection of interaction demanding and nodding assent based on stereo vision," *J. Phys. Agents*, vol. 1, no. 1, pp. 15–26, 2007.

[57] A. Fern'ndez-Baena, A. Susin, and X. Lligadas, "Biomechanical validation of upper-body and lower-body joint movements of Kinect motion capture data for rehabilitation treatments," in *Proc. 4th Int. Conf. Intell. Netw. Collaborative Syst.*, Sep. 2012, pp. 656–661.

[58] M. L. Walters, D. S. Syrdal, K. L. Koay, K. Dautenhahn, and R. Te Boekhorst, "Human approach distances to a mechanical-looking robot with different robot voice styles," in *Proc. 17th IEEE Int. Symp. Robot Hum. Interact. Commun. (RO-MAN)*, Aug. 2008, pp. 707–712.

[59] P. A. Hancock, D. R. Billings, K. E. Schaefer, J. Y. Chen, E. J. De Visser, and R. Parasuraman, "A meta-analysis of factors affecting trust in human-robot interaction," *Hum. Factors*, vol. 53, no. 5, pp. 517–527, 2011.

[60] K. Pitsch, H. Kuzuoka, Y. Suzuki, L. Sussenbach, P. Luff, and C. Heath, "'The first five seconds': Contingent stepwise entry into an interaction as a means to secure sustained engagement in HRI," in *Proc. 18th IEEE Int. Symp. Robot Hum. Interact. Commun. (RO-MAN)*, Sep. 2009, pp. 985–991.

[61] E. Law, V. Cai, Q. F. Liu, S. Sasy, J. Goh, A. Blidaru, and D. Kulić, "A wizard-of-oz study of curiosity in human-robot interaction," in *Proc. 26th IEEE Int. Symp. Robot Hum. Interact. Commun. (RO-MAN)*, Aug. 2017, pp. 607–614.

[62] D. Hebesberger, T. Koertner, C. Gisinger, and J. Pripfl, "A long-term autonomous robot at a care hospital: A mixed methods study on social acceptance and experiences of staff and older adults," *Int. J. Social Robot.*, vol. 9, no. 3, pp. 417–429, 2017.

[63] J. Miehle, I. Bagci, W. Minker, and S. Ultes, "A social companion and conversational partner for the elderly," in *Advanced Social Interaction With Agents*. Springer, 2017, pp. 103–109.

[64] T. Kanda, M. Shiomi, Z. Miyashita, H. Ishiguro, and N. Hagita, "An affective guide robot in a shopping mall," in *Proc. 4th ACM/IEEE Int. Conf. Hum.-Robot Interact. (HRI)*, Mar. 2009, pp. 173–180.

[65] Y. Iwamura, M. Shiomi, T. Kanda, H. Ishiguro, and N. Hagita, "Do elderly people prefer a conversational humanoid as a shopping assistant partner in supermarkets?" in *Proc. 6th ACM/IEEE Int. Conf. Hum.-Robot Interact. (HRI)*, Mar. 2011, pp. 449–457.

[66] Z. Zhang, "Microsoft Kinect sensor and its effect," *IEEE MultiMedia*, vol. 19, no. 2, pp. 4–10, Feb. 2012.

[67] M. A. Yousuf, Y. Kobayashi, Y. Kuno, A. Yamazaki, and K. Yamazaki, "How to move towards visitors: A model for museum guide robots to initiate conversation," in *Proc. IEEE RO-MAN*, Aug. 2013, pp. 587–592.

[68] E. Grigore, A. Roncone, O. Mangin, and B. Scassellati, "Preference-based assistance prediction for human-robot collaboration tasks," in *Proc. Intell. Robots Syst. (IROS)*, Oct. 2018, pp. 4441–4448.

[69] E. C. Grigore, O. Mangin, A. Roncone, and B. Scassellati, "Predicting supportive behaviors for human-robot collaboration," in *Proc. 17th Int. Conf. Auto. Agents MultiAgent Syst.*, 2018, pp. 2186–2188.

[70] T. Stipancic, B. Jerbic, and P. Curkovic, "A context-aware approach in realization of socially intelligent industrial robots," *Robot. Comput.-Integr. Manuf.*, vol. 37, pp. 79–89, Feb. 2016.

[71] A. M. von der Pütten, N. C. Krämer, C. Becker-Asano, K. Ogawa, S. Nishio, and H. Ishiguro, "At the café—Exploration and analysis of people's nonverbal behavior toward an android," in *Geminoid Studies: Science and Technologies for Humanlike Teleoperated Androids*. Springer, 2018, pp. 375–397.

[72] S. Andrist, B. Mutlu, and A. Tapus, "Look like me: Matching robot personality via gaze to increase motivation," in *Proc. 33rd Annu. ACM Conf. Hum. Factors Comput. Syst.*, 2015, pp. 3603–3612.

[73] J.-S. Hu, J.-J. Wang, and D. M. Ho, "Design of sensing system and anticipative behavior for human following of mobile robots," *IEEE Trans. Ind. Electron.*, vol. 61, no. 4, pp. 1916–1927, Apr. 2014.

[74] C. Liang, L. Qi, and L. Guan, "Motion energy guided multi-scale heterogeneous features for 3D action recognition," in *Proc. IEEE Vis. Commun. Image Process. (VCIP)*, Dec. 2017, pp. 1–4.

[75] M. Moradi, M. Moradi, and F. Bayat, "On robot acceptance and adoption a case study," in *Proc. 8th Conf. AI Robot. 10th RoboCup Iranopen Int. Symp. (IRANOPEN)*, Apr. 2018, pp. 21–25.

[76] N. Mavridis, "A review of verbal and non-verbal human–robot interactive communication," *Robot. Auto. Syst.*, vol. 63, pp. 22–35, Jan. 2015.

[77] B. van den Brink and Z. Yumak, "Social gaze model for an interactive virtual character," in *Proc. Int. Conf. Intell. Virtual Agents*. Springer, 2017, pp. 451–454.

[78] H. P. C. Sirithunge, A. G. B. P. Jayasekara, and D. P. Chandima, "Human attention estimation by a domestic service robot using upper body skeletal information," in *Proc. IEEE Int. Conf. Ind. Inf. Syst. (ICIIS)*, Dec. 2017, pp. 1–6.

[79] H. P. C. Sirithunge, M. A. V. J. Muthugala, A. G. B. P. Jayasekara, and D. P. Chandima, "Interpretation of interaction demanding of a user based on nonverbal behavior in a domestic environment," in *Proc. IEEE Int. Conf. Fuzzy Syst. (FUZZ-IEEE)*, Jul. 2017, pp. 1–8.

[80] M. Bennewitz, F. Faber, D. Joho, M. Schreiber, and S. Behnke, "Towards a humanoid museum guide robot that interacts with multiple persons," in *Proc. 5th IEEE-RAS Int. Conf. Humanoid Robots*, Dec. 2005, pp. 418–423.

[81] R. Mead and M. J. Matarić, "Proxemics and performance: Subjective human evaluations of autonomous sociable robot distance and social signal understanding," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Sep. 2015, pp. 5984–5991.

[82] R. Mead and M. J. Matarić, "Autonomous human–robot proxemics: Socially aware navigation based on interaction potential," *Auto. Robots*, vol. 41, no. 5, pp. 1189–1201, 2017.

[83] M. Strait, C. Canning, and M. Scheutz, "Let me tell you! Investigating the effects of robot communication strategies in advice-giving situations based on robot appearance, interaction modality and distance," in *Proc. ACM/IEEE Int. Conf. Hum.-Robot Interact.*, 2014, pp. 479–486.

[84] A. Birk, "Robot learning and self-sufficiency: What the energy-level can tell us about a robot's performance," in *Proc. Eur. Workshop Learn. Robots*. Springer, 1997, pp. 109–125.

[85] J. Peltason and B. Wrede, "Pamini: A framework for assembling mixed-initiative human-robot interaction from generic interaction patterns," in *Proc. 11th Annu. Meeting Special Interest Group Discourse Dialogue*, 2010, pp. 229–232.

[86] A. Bauer, K. Klasing, G. Lidoris, Q. Mühlbauer, F. Rohrmüller, S. Sosnowski, T. Xu, K. Kühnlenz, D. Wollherr, and M. Buss, "The autonomous city explorer: Towards natural human-robot interaction in urban environments," *Int. J. Social Robot.*, vol. 1, no. 2, pp. 127–140, 2009.

[87] M. L. Spekman, E. A. Konijn, and J. F. Hoorn, "Perceptions of healthcare robots as a function of emotion-based coping: The importance of coping appraisals and coping strategies," *Comput. Hum. Behav.*, vol. 85, pp. 308–318, Aug. 2018.

[88] P. H. Kahn, Jr., H. Ishiguro, B. Friedman, and T. Kanda, "What is a human?: Toward psychological benchmarks in the field of human–robot interaction," in *Proc. 15th IEEE Int. Symp. Robot Hum. Interact. Commun. (ROMAN)*, Sep. 2006, pp. 364–371.

[89] M. L. Walters, D. S. Syrdal, K. Dautenhahn, R. Te Boekhorst, and K. L. Koay, "Avoiding the uncanny valley: Robot appearance, personality and consistency of behavior in an attention-seeking home scenario for a robot companion," *Auto. Robots*, vol. 24, no. 2, pp. 159–178, 2008.

[90] C. Rich, B. Ponsler, A. Holroyd, and C. L. Sidner, "Recognizing engagement in human-robot interaction," in *Proc. 5th ACM/IEEE Int. Conf. Hum.-Robot Interact. (HRI)*, Mar. 2010, pp. 375–382.

[91] E. Martinson and D. Brock, "Improving human-robot interaction through adaptation to the auditory scene," in *Proc. 2nd ACM/IEEE Int. Conf. Hum.-Robot Interact. (HRI)*, Mar. 2007, pp. 113–120.

[92] O. A. I. Ramírez, H. Khambhaita, R. Chatila, M. Chetouani, and R. Alami, "Robots learning how and where to approach people," in *Proc. 25th IEEE Int. Symp. Robot Hum. Interact. Commun. (RO-MAN)*, Aug. 2016, pp. 347–353.

[93] K. Fischer, L. C. Jensen, S.-D. Suvei, and L. Bodenhagen, "Between legibility and contact: The role of gaze in robot approach," in *Proc. 25th IEEE Int. Symp. Robot Hum. Interact. Commun. (RO-MAN)*, Aug. 2016, pp. 646–651.

[94] H. M. R. T. Bandara, M. A. V. J. Muthugala, A. G. B. P. Jayasekara, and D. P. Chandima, "Cognitive spatial representative map for interactive conversational model of service robot," in *Proc. 27th IEEE Int. Symp. Robot Hum. Interact. Commun. (RO-MAN)*, Aug. 2018, pp. 686–691.

[95] J. Baraglia, M. Cakmak, Y. Nagai, R. P. Rao, and M. Asada, "Efficient human-robot collaboration: When should a robot take initiative?" *Int. J. Robot. Res.*, vol. 36, nos. 5–7, pp. 563–579, 2017.

[96] S. Oishi, Y. Kohari, and J. Miura, "Toward a robotic attendant adaptively behaving according to human state," in *Proc. 25th IEEE Int. Symp. Robot Human Interact. Commun. (RO-MAN)*, Aug. 2016, pp. 1038–1043.

[97] Y. Gao, S. Wallkötter, M. Obaid, and G. Castellano, "Investigating deep learning approaches for human-robot proxemics," in *Proc. 27th IEEE Int. Symp. Robot Hum. Interact. Commun. (RO-MAN)*, Aug. 2018, pp. 1093–1098.

[98] A. Cruz-Maya and A. Tapus, "Adapting robot behavior using regulatory focus theory, user physiological state and task-performance information," in *Proc. 27th IEEE Int. Symp. Robot Hum. Interact. Commun. (RO-MAN)*, Aug. 2018, pp. 644–651.

[99] J. James, C. I. Watson, and B. MacDonald, "Artificial empathy in social robots: An analysis of emotions in speech," in *Proc. 27th IEEE Int. Symp. Robot Hum. Interact. Commun. (RO-MAN)*, Aug. 2018, pp. 632–637.

[100] E. T. Hall, R. L. Birdwhistell, B. Bock, P. Bohannan, A. R. Diebold, Jr., M. Durbin, M. S. Edmonson, J. L. Fischer, D. Hymes, S. T. Kimball, W. La Barre, F. Lynch, J. E. McClellan, D. S. Marshall, G. B. Milner, H. B. Sarles, G. L. Trager, and A. P. Vayda, "Proxemics [and comments and replies]," *Current Anthropol.*, vol. 9, nos. 2–3, pp. 83–108, 1968.

[101] A. Kendon, *Conducting Interaction: Patterns of Behavior in Focused Encounters*, vol. 7. Cambridge, U.K.: Cambridge Univ. Press, 1990.

[102] J. Sanghvi, G. Castellano, I. Leite, A. Pereira, P. W. McOwan, and A. Paiva, "Automatic analysis of affective postures and body motion to detect engagement with a game companion," in *Proc. 6th Int. Conf. Hum.-Robot Interact.*, 2011, pp. 305–312.

[103] C. Shi, S. Satake, T. Kanda, and H. Ishiguro, "A robot that distributes flyers to pedestrians in a shopping mall," *Int. J. Social Robot.*, vol. 10, no. 4, pp. 421–437, 2018.

[104] H. Akaike, "Fitting autoregressive models for prediction," *Ann. Inst. Stat. Math.*, vol. 21, no. 1, pp. 243–247, Dec. 1969.

[105] Y. Mohammad, T. Nishida, and S. Okada, "Unsupervised simultaneous learning of gestures, actions and their associations for human-robot interaction," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Oct. 2009, pp. 2537–2544.

CHAPA SIRITHUNGE (S'17) received the B.Sc. degree in engineering from the University of Moratuwa, Sri Lanka, in 2016. She is currently a Postgraduate Student with the Robotics and Control Laboratory, Department of Electrical Engineering, University of Moratuwa. She is working as a Research Assistant with the University of Moratuwa, from 2016. Her research interests include cognitive robotics, intelligent systems, machine learning, and human–robot interaction.

A. G. BUDDHIKA P. JAYASEKARA (M'07–SM'17) received the B.Sc. and M.Sc. degrees in electrical engineering from the University of Moratuwa, respectively, in 2004 and 2006, and the doctoral degree in robotics and intelligent systems from Saga University, Japan. He received the Japanese Government Monbukagakusho Scholarship to pursue postgraduate studies in Japan. He joined the Department of Electrical Engineering, University of Moratuwa, after completing his doctoral studies in 2010, where he currently a Senior Lecturer.

His main research interests include artificial intelligence techniques, human–robot interaction, human-friendly robotics, machine learning, and intelligent systems. He is the Principle Investigator of the Intelligent Service Robotics Group, Department of Electrical Engineering, University of Moratuwa, and working as the Director of Center for Advanced Robotics (CAR), a multidisciplinary research center of University of Moratuwa from 2017. Currently, he is working on two main projects: intelligent mobile robot for domestic environments (MIRob-Moratuwa Intelligent Robot) and intelligent wheelchair for domestic environments.

D. P. CHANDIMA (M'03–SM'17) received the B.Sc. degree from the University of Moratuwa, Sri Lanka, in 2001, and the M.Eng. degree in advanced systems control engineering and Ph.D. degree in robotics and intelligent systems, in 2006 and 2009, respectively, from University of Saga, Japan.

Currently, he is a Senior Lecturer with the Department of Electrical Engineering, University of Moratuwa, Sri Lanka. His research interests include estimation theoretic mobile robot navigation, multi-target tracking, data association, power electronic converters, control of microgrid systems, energy storage integrations, and electrical machines.