Title	Robot Social Emotional Development through Memory Retrieval
Author(s)	Bui, Ha-Duong; Dang, Thi Le Quyen; Chong, Nak Young
Citation	2019 7th International Conference on Robot Intelligence Technology and Applications (RiTA): 46-51
Issue Date	2019-11
Type	Conference Paper
Text version	author
URL	http://hdl.handle.net/10119/16206
Rights	This is the author's version of the work. Copyright (C) 2019 IEEE. 2019 7th International Conference on Robot Intelligence Technology and Applications (RiTA), 2019, 46-51. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.
Description	

Robot Social Emotional Development through Memory Retrieval*

Ha-Duong Bui[1], Thi Le Quyen Dang[1], and Nak Young Chong[1]

Abstract— **Robot emotion representation is gaining increasing attention to facilitate long-term human-robot interaction (HRI) in recent years. In particular, human-like robot emotion elicited through HRI is of great use in creating trust between humans and robots. In attempting to represent robot emotions that lead to gaining social acceptance, psychological studies of human emotion have been extensively performed. Among the various factors that affect the way people express their emotional competencies, we conjecture that two factors, social interaction and experience, can be considered important to elicit human emotions, and therefore can be used to represent robot emotions. We believe that social and developmental interaction paradigms, such as social sharing and social referencing, can shape robot emotions toward promoting social acceptance. Besides, the robot's previous experience can be a key factor contributing to robot personality formation and development. In this paper, we not only focus on the modeling of two eliciting factors affecting the formation of robot emotion but also examine the decline of memory retention over time. Specifically, the relationship between emotion and memory is investigated to design a filter for the memory consolidation process and memory forgetting mechanism. The mechanism is used to enhance robot memory performance based on emotional salience and time parameters. Experiments were performed with a humanoid robot Pepper having verbal and non-verbal interactions with 24 human subjects. Participants rate their perception of the robot in terms of human-likeness, likeability, safety, and emotional expressions through a questionnaire. The results showed that most of the participants enjoyed interacting with the robot and they wished they could have more interactions in the future. They perceived safety and responded favorably toward the robot emotional expressions.**

I. INTRODUCTION

Emotion helps reflect how an individual is affected by and adapted to environmental stimuli or situation during human-human interaction, thereby it helps create a bond or trust between humans. When interacting with a social robot, humans tend to treat the robot as a companion/friend/pet. Expecting the same effect of emotion in human-human interaction, artificial emotions of robots have been studied to boost human-robot interaction (HRI) [1]. Likewise, robot emotions should be generated and expressed through behaviors appropriately to gain human acceptance. Furthermore, robot emotions have been modeled as a function of cognitive memory to recall affective experiences when selecting a specific behavior, inspired by the role of human emotion in reliving past experiences [2], [3]. However, it should be noted that human emotions are influenced by past experiences through the retrieval process of stored memories, which has been neglected in previous literature. Besides, human emotions are activated and influenced by other factors known as the social effects which consist of two aspects: referencing and sharing. On the one hand, social referencing helps human infants acquire basic interpretation to generate emotions and imitate their parents' emotional behaviors. On the other hand, social sharing helps an individual gain a more detailed knowledge and drive their emotions based on personal knowledge and experiences during self-discovery and exploration of the environment.

In this paper, we propose a new robot emotion generation architecture based on robot personal experiences and human guide reflecting the effects of social referencing and social sharing. To enable robots to consolidate, maintain, and recall personal experience, we develop our robot long-term memory based on the Epigenetic Robot Intelligent System (ERIS) [4]. In addition, human guides during HRI help direct and shape robot emotions to promote human acceptance of social robots. Our robot emotion is represented on the valence-arousal space [5] which was used in modeling emotions in cognitive science [6], [7]. Besides, we used emotional body expressions proposed by our previous work [8], [9] to enable robots to show their emotional states through their posture and gestures. In this work, to evaluate the human understanding and acceptance of robots with the proposed social emotional development architecture, we employed a humanoid robot Pepper [10] and performed verbal and non-verbal interactions between humans and the robot. In our experiments, human participants were required to help Pepper acquire interpretation about an object presented to generate and express emotions toward the object. Then the Pepper's responses were evaluated by the participants.

This work makes a contribution to the ongoing Horizon 2020 EU-Japan project "Culture-Aware Robots and Environmental Sensor Systems for Elderly Support" (CARESSES) [11], [12], which aims to pave the way toward integrating the use of care robots and smart environments [13], [14] into the future of societal infrastructure.

This paper is organized as follows: In Section II, we present our proposed method for representing robot emotions based on the social effects and robot personal experiences. Experimental scenarios, procedures, results are described in Section III. And our conclusions are drawn in Section IV.

*This work was supported by the EU-Japan coordinated R&D project on "Culture Aware Robots and Environmental Sensor Systems for Elderly Support," commissioned by the Ministry of Internal Affairs and Communications of Japan and EC Horizon 2020 Research and Innovation Programme under grant agreement No. 737858. The authors are also grateful for financial supports from the Air Force Office of Scientific Research (AFOSR-AOARD/FA2386-19-1-4015).

[1]Ha-Duong Bui, Thi Le Quyen Dang, and Nak Young Chong are with the School of Information Science, Japan Advanced Institute of Science and Technology, 1-1 Asahidai, Nomi, Ishikawa 923-1292, Japan {bhduong, quyen.dang, nakyoung}@jaist.ac.jp

II. Representing Robot Emotion based on Social Effects and Personal Experience

To enable natural interactions between a human and a robot and offer deeper levels of engagement for a long period of time, we propose a novel robot emotion generation model incorporating the robot's personal experience and social effects for both perspectives of referencing and sharing. The details of the proposed method are given in the following sections.

A. ERIS-based Memory Architecture

Cognitive memory plays an important role in the development of social robots for active and autonomous learning responding to environmental stimuli through acquired experiences and continuous interactions. Among different memory architectures, ERIS [4] aims at modeling interconnections among memory components: Semantic Memory (SM), Episodic Memory (EM) and Procedural Memory (PM). This architecture was used to develop a robot long-term memory (LTM) for robot emotional experiences in our previous work [7]. Extending the previous work, we now focus on the filtering process in memory consolidation and the effect of memory forgetting for removing out-of-date information and facilitating memory performance. Besides, we also design a simple PM component to enable robots to maintain predefined emotional expressions and behaviors.

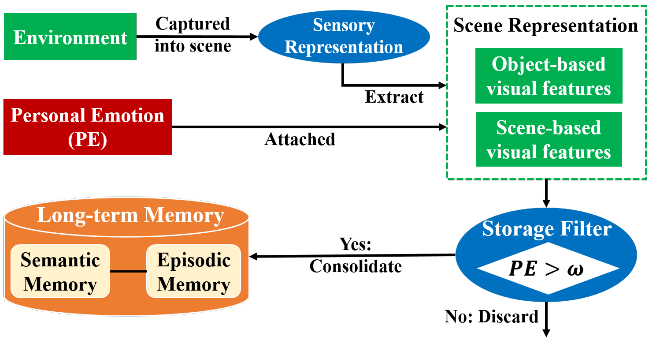


Fig. 1. Memory Consolidation Process

1) Memory Consolidation: The memory architecture helps a robot consolidate its knowledge related to the surrounding environment and objects therein. The consolidation process is shown in Fig. 1. To extract visual features from the scene including objects for forming affective appraisals for future access, we design a functional module called Sensory Representation (SR) which can perform both global extraction and local extraction. Global extraction can obtain scene-based visual features such as the number of objects and positions of objects in the scene. Local extraction helps get object-based visual features such as color or SIFT features. Extracted features from SR are combined with the updated robot personal emotion responding to the scene to be encoded as an experience. We use a predefined threshold ω indicating the minimum level of influence that must be exerted to the memory consolidation process. If the emotion at the time of consolidating the experience (PE) is weaker than ω, the current experience will be discarded; otherwise, it can be consolidated into the robot LTM with a certain identification. Scene-based visual features combined with the updated personal emotion can be consolidated as an episodic memory item. Object-based visual features can form semantic memory items. One semantic item can represent features of only one object. One experience is represented by only one episodic memory item and one or more semantic memory items. Those memory items are linked to represent the interconnection among memory components in the robot LTM. Those links can help the robot relive past experiences when they are recalled through memory retrieval.

2) Memory Retrieval: When observing a scene which contains a single object, robots are able to recall past experiences based on object-based visual features [6]. They extracted features of the new object (o_n) using the SR function to compare with the corresponding features of the previously seen objects (o_s) to find the most similar object. The similarity level of two different objects are defined using the object color histogram and SIFT features given by Eq. 1

$$sim(o_n, o_s) = \varepsilon \cdot sim_{cl}(o_n, o_s) + (1-\varepsilon) \cdot sim_{sift}(o_n, o_s) \quad (1)$$

where sim_{cl} and sim_{sift} are calculated based on the Euclidean distance between the color histograms and SIFT features of two objects, respectively. ε represents the influence proportion of different features. We define a threshold η and compare η with the similarity scores of all seen objects and the new object. If $sim(o_n, o_{s_i}) > \eta$, the i^{th} seen object is similar to the new object. If multiple similar objects exist, the object with the highest similarity score is selected as the most similar previously seen object to the new object. The selected object has been consolidated as a semantic item which is linked to a certain episodic memory item. Recalling these memory items enables the robot to relive the past experience including visual information and emotion. In our design, we assume that the memory forgetting mechanism, mood influence, and memory recollection faults do not hinder the retrieval of memories.

3) Memory Forgetting Mechanism: In an attempt to design an effective forgetting mechanism, the time decay-based forgetting mechanism was applied to a previous memory architecture as an emotion-based function calculating the historical component of episodic memory items [15]. The historical component was combined with a context component to decide whether a memory item should be retrieved by a given object or not. We use this historical component to identify the active level of each memory item to remove out-of-date information. Besides, both of semantic and episodic items contain declarative knowledge which can be forgotten after a period of time equally. The active level is used to identify only episodic items, because the emotional information of an experience is consolidated in an episodic memory item and we consider the emotional salience as an important factor to activate the memory item for the retrieval process. Semantic items can be erased when all linked episodic items are deleted. The active level of an episodic item (E) at time

t is calculated by

$$ActiveLevel_E(t) = \frac{v+n}{M(t)+b} \cdot s(t) \quad (2)$$

where v and b are set to 1 and 3, respectively, for shaping the decay curve [15], n is the number of retrieval times, $s(t)$ is the decay function with emotional salience and $M(t)$ is the integral of $s(t)$. Specifically, the decay function is calculated by

$$s(t) = \mathrm{e}^{-\alpha \cdot t} \quad (3)$$

where t is the lifetime from the instant the item is created until the current time measured in minutes and α is the emotional salience. We apply the arousal value (a_{value}) of the emotion attached to each episodic memory item, since arousal can influence the effect of valence on memory performance [16], to calculate α given by

$$\alpha = \frac{(a_{max} - a_{value})}{a_{max}} \quad (4)$$

where a_{max} is the maximum value of arousal component in a certain range. If a memory item is consolidated with arousal component that equals a_{max}, the item is always available for retrieval. We normalize the active level of each memory item at the current time t_c in the range of (0,1) to get the retrieval likelihood of each memory item as follows:

$$Likelihood_E(t_c) = \frac{ActiveLevel_E(t_c)}{\sum_{j=1}^{k} ActiveLevel_{E_j}(t_c)} \quad (5)$$

where $ActiveLevel_E(t_c)$ and $ActiveLevel_{E_j}(t_c)$ are the history components of the current memory item and the j^{th} memory item at the current time, respectively.

After getting the retrieval possibility of an episodic memory item $Likelihood_E$, we compare it with a threshold fg_{thresh} to decide whether the memory item should be kept or removed. If $Likelihood_E \leq fg_{thresh}$, then the memory item should be removed. We do not apply the mechanism for pre-defined behaviors maintained in PM.

B. Robot Emotion Generation Model

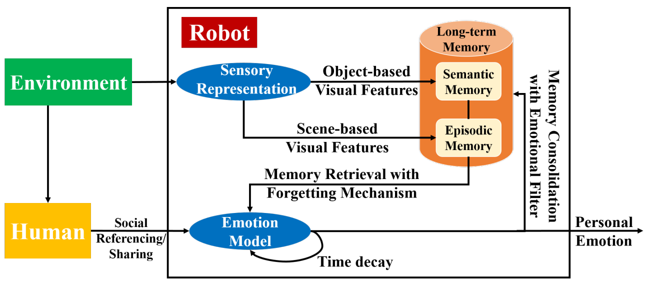


Fig. 2. Robot Emotion Generation Model

Our robot emotion model aims at incorporating two main factors capable of enabling the robot to generate and express its internal states in an interactive and developmental manner. The first factor is the social effects modeled as social sharing and social referencing. The robot's personal emotion and interpretation about the environment are learned and affected by humans during HRI. The second factor is the robot's own personal experiences acquired continuously throughout the robot's developmental process. These experiences can help the robot create its evolving personality, as well as differentiate itself from other robots. Furthermore, we also consider the decay of the intensity of the emotion over time and devise a detailed mechanism to optimize the memory performance. The general diagram of our system is shown in Fig. 2.

During interactions with a human, the robot receives visual stimuli from the environment and captures scenes. The robot is endowed with the Sensory Representation (SR) function capable of extracting visual features from the scene to be ready for memory processes. Affective appraisals, which are created based on object-based visual features extracted through SR, are used to search for the most similar previously seen object through the memory retrieval process in Section II-A.2. The experience that contains the most similar object is recalled from the robot LTM. Successfully recalling the most similar experience which contains the seen object, the robot can relive the experience including all sensory information and emotion. This experience can influence the robot personal emotion. Based on the assumption made in [7], only experienced emotion can influence the robot personal emotion. Furthermore, we ignore the application of forgetting mechanisms on decaying the experienced emotion based on the assumption in Section II-A.3.

In addition, the human may inject his/her emotion into the robot. This is what we call the guided emotion component directing and influencing the robot personal emotion. In this work, we enable the robot to interact with humans through spoken natural language as the same way a human interacts with others. Guidance from humans can be given through utterances which describe their emotions and abstract thought about the given object such as *It looks scary* or *I am afraid of it*. Since we focus on modeling and representing robot emotions, we ignore complicated natural language processing techniques. In our implementation, we use a maximum of three sentences to describe the user's guided emotion directed to an object. The robot is equipped with the Hint Extraction function which can convert speech to text, split the sentence into words and obtain guided emotion based on split words. In order to get guided emotions from words, we consider the word frequency to choose the most suitable word for describing human emotions. Low-frequency words require higher brain effort to recognize [17], while emotions and emotional effects decay over time [18]. Thus, high-frequency words can be used to express emotions without changing the emotions through time effects. The Twister Data Server and an extensive list of English words with affective ratings [19] are used as a frequency dictionary and an emotional dictionary, respectively. On the one hand, the frequency dictionary contains multiple words with different popularity in daily use. On the other hand, the emotion dictionary contains 13,915 words, each of which was rated by human subjects to have values of valence and arousal with a standard mean and deviation. Referring to the frequency dictionary, the robot can acquire most popular words split from human guide. Each word gives the robot different values of valence and arousal through a searching mechanism applied on the emotion dictionary. Words with no values of valence and arousal will not influence the robot, thus we assign frequency values of these words as 0. We calculate the guided emotion based on the rated emotional values of selected words and the word frequency as follows:

$$C_{GE} = \frac{\sum_i^m Cword_i \cdot freq_i}{\sum_i^m freq_i} \tag{6}$$

where m is the number of the most important and popular words, $Cword_i$ represents the emotional value of the i^{th} word taken from the emotion dictionary and $freq_i$ is the number of times of occurrence of the i^{th} word in the frequency dictionary.

Now the robot's personal emotion that consists of the human guided emotion and the experienced emotion is updated as follows [7]:

$$C_{PE_{t_2}} = \lambda_1 \cdot C_{GE_{t_1}} + \lambda_2 \cdot C_{EE_{t_1}} + (1 - \lambda_1 - \lambda_2) \cdot C_{PE_{t_1}} \tag{7}$$

where $C_{PE_{t_2}}$ describes the current emotion of the robot. $C_{GE_{t_1}}$, $C_{EE_{t_1}}$, and $C_{PE_{t_1}}$ represent the guided emotion, the experienced emotion, and the robot emotion at the previous time step, respectively. λ_1 and λ_2 are percentage weights for $C_{GE_{t_1}}$ and $C_{EE_{t_1}}$, respectively, based on their relative importance. Since valence and arousal act differently on directing robot attentions toward humans and memory performance enhancement, the parameter μ is used as the least influence of each component to calculate λ_1 and λ_2 as follows:

$$\lambda_1 = \frac{|C_{GE_{t_1}} - \mu|}{|C_{GE_{t_1}} - \mu| + |C_{EE_{t_1}} - \mu| + |C_{PE_{t_1}} - \mu|} \tag{8}$$

$$\lambda_2 = \frac{|C_{EE_{t_1}} - \mu|}{|C_{GE_{t_1}} - \mu| + |C_{EE_{t_1}} - \mu| + |C_{PE_{t_1}} - \mu|} \tag{9}$$

High arousal can influence the human tendency to share with others and enhance memory performance. Negative and positive valences do the similar effects. Therefore, we set μ to 1 and 5, respectively, to update the valence and arousal components of the robot personal emotion [7].

The robot personal emotion is used not only as a filter for the memory consolidation process, but also as a decisive factor in selecting emotional behaviors. We enable the robot to express emotions through its posture and gesture, thereby humans can understand how the robot is affected by and adapted to a variety of environmental stimuli. Emotional expressions can help increase the human comfort level when interacting with robots. A basic behavioral repertoire using eight emotional labels as shown in Fig. 3 is implemented on the Pepper robot [8], with three different movement speeds: fast, medium, and slow. Different combinations of valence and arousal values are shown in Fig. 4. Incorporating different movement speeds into a basic repertoire, we can endow the Pepper robot with 24 different emotional body expressions.

III. Experiments

A. Participants

A total of 24 participants (12 males, 12 females) ranging in age from 22 to 35 ($M = 26.42, SD = 3.8$) took part in the experiment. They are from Vietnam, China, Thailand, Malaysia, Japan, Mongolia, Taiwan, Syria, and Indonesia and recruited at Japan Advanced Institute of Science and Technology (JAIST). All participants are students in JAIST who are enrolled in a Master's or Doctoral program in English. Before starting the actual experiments, all participants were given a written instruction about the experimental protocol. They could ask any questions until they can comprehend the protocol.

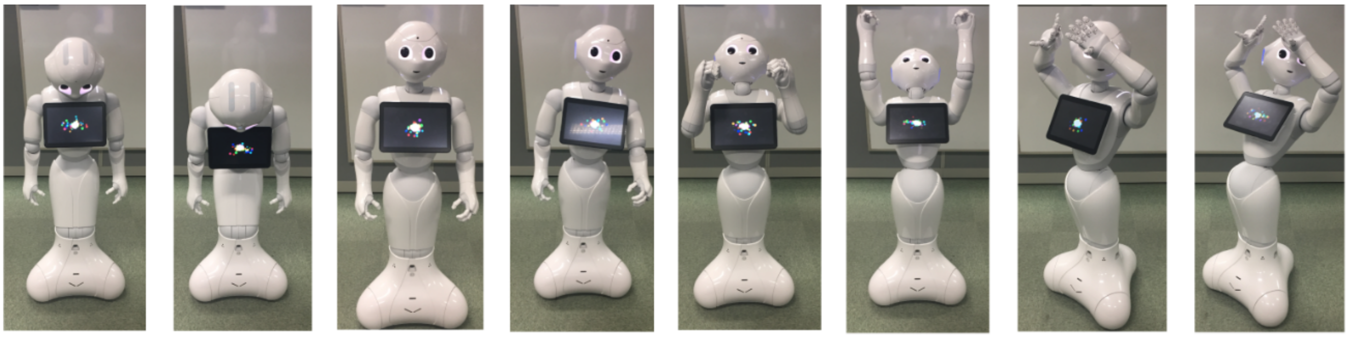
B. Experiment Scenario

We have designed an experimental scenario where the Pepper robot interacts with each of the participants to develop its social emotion and behavioral expressions. The scenario employs two phases of processing: training and testing.

The two phase experiment investigated that the robot can create object affective appraisals based on the visual features underlying object recognition such as color and shape, recall past experiences, and generate personal emotions and expressions given each interaction with the participant. During interactions, the environmental scene and a single object therein were captured by the robot's RGB camera. The robot emotion was initialized by a neutral non-arousing state before starting both phases. The robot LTM is empty at the beginning of the training phase. In addition, we prohibited the robot from interacting with objects and/or the participants between the two phases to let the robot emotion return to a neutral non-arousing state.

The objective of the first phase (training phase) is to enable the robot to acquire knowledge and shape emotions to increase acceptance to the participant. In this phase, the robot is directed and influenced by the human participant's emotional guide. The guide is given to the robot through the utterances of participants. By doing so, each of the participants' emotion is injected to the robot. The injected or guided emotion is then used as one of the factors contributing to the robot personal emotion update. Besides, the visual features of the objects are extracted to form affective appraisals. These appraisals are used to recall the robot's past experiences through the memory retrieval process, thereby the experienced emotion is retrieved. The robot personal emotion is updated based on the human-guided emotion and the experienced emotion to select a behavioral expression. Finally, the participant feedback survey is administered through a questionnaire.

The second phase (testing phase) aims to test whether the robot can develop and express its emotion using the human guides injected in the first phase. There are no hints given to the robot in this phase. The robot acquires the human knowledge, forms affective appraisals, and performs reasoning to recall its past experiences. The experienced emotions are used as the only factor contributing to the



(a) somewhat sad (b) sad (c) somewhat relaxed (d) relaxed (e) somewhat happy (f) happy (g) somewhat afraid (h) afraid

Fig. 3. Robot behavioral repertoire based on eight basic emotion labels [8]

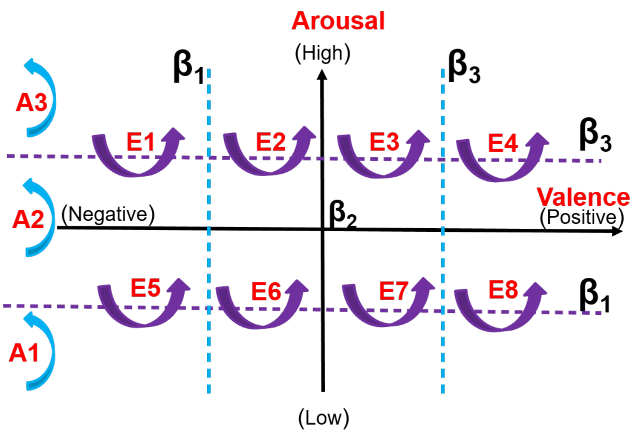


Fig. 4. Mapping robot emotions to body expressions. $E_1, E_2, ..., E_8$ represent basic behavioral repertoire (Fig. 3). A_1, A_2, A_3 show the speed of behaviors. β_1, β_2, β_3 are the range parameters to classify different regions in the valence-arousal space [8].

robot's personal emotion update responding to the visual stimuli. As was the case with the first phase, the robot behavioral expressions are observed and evaluated by the participant.

At the end of each phase, participants were asked to fill out questionnaires written specifically for their perception and acceptance of the robot. We evaluate three main characteristics of the robot such as *human-likeness* [20], *likeability*, and *safety* [21], as well as the robot emotional body expressions with 5-point Likert type scales (strongly disagree = 1; strongly agree = 5). The survey items are shown in Table I.

C. Results and Discussion

We perform the ANOVA test to get the confident interval for pairwise comparisons between the mean values of each testing item. The mean values and standard deviation of item details and the p-value of all items are shown in Table I. Participants confirm that the robot emotional expressions and appearance influence their evaluation on the robot's *human-likeness*. This reflects their attribution of human nature traits to the robot (*F(2.70) = 4.21, p = 0.008*). In addition, those expressions and the robot appearance also affect participants' perception of the robot's *likeability* (*F(3.13) = 3.9, p = 0.024*). According to the data, the participants' perception of *human-likeness* depends highly on the "friendly" item compared to other items. They positively think that the robot has *human-likeness* characteristics especially in terms of "friendly" and "high cognitive ability" items. Furthermore, in the *likeability* index, emotional responses of the robot during HRI pleased the participants. Most participants wished to have future interactions with the robot ($M = 4.3, SD = 0.73$), since they enjoyed interacting with the robot ($M = 4, SD = 0.83$). Regarding the perception of *safety*, most participants do not feel "anxious" ($M = 2.14, SD = 1.02$) or "scary" ($M = 2.08, SD = 0.92$) toward the robot. They show their favor when having interactions with the robot not only based on the robot *human-likeness* but also for the responses of the robot toward given objects and human guide. Thus, the *likeability* characteristics got a higher rating ($M = 3.94, SD = 0.71$) compared to *human-likeness*. Besides, they also do not feel that the robot is "aggressive" ($M = 1.85, SD = 0.99$). Participants also show their satisfaction in response to the robot movement ($M = 3.46, SD = 0.61$) and emotional body expressions ($M = 3.62, SD = 0.78$). The low rating of factors ($M = 1.94, SD = 0.91$) used to measure the participants' perception of *safety* implied that they did not feel "anxious", "scary" or "aggressive".

IV. CONCLUSION

In this paper, we have proposed a new developmental approach to the construction and use of robot emotion generation architecture to fulfill the user expectations and acceptance of social robots. The architecture fully incorporated two main factors contributing to the enhancement of robot emotions: (1) social effects such as social referencing and social sharing and (2) robot personal experiences. We have designed the robot long-term memory by extending the developmental memory architecture ERIS whereby the robot could acquire knowledge, obtain emotional experiences, and form affective appraisals for future access. A memory forgetting mechanism was developed based on emotional salience and time parameters to remove out-of-date information and improve the memory performance. Furthermore, the robot personal emotion was modeled based on a decay function

TABLE I

QUESTIONNAIRE ITEMS FOR EVALUATING PARTICIPANTS' PERCEPTION OF THE ROBOT (5 - STRONGLY AGREE; 1 - STRONGLY DISAGREE) AND DESCRIPTIVE STATISTICS OF PARTICIPANTS' RATING ($^{***}p < 0.01$ AND $^{**}p < 0.05$)

Evaluation point	Details	Mean	SD	*p-value*
Human-likeness				
	1. Friendly	4.08	0.67	
	2. Believable	3.46	0.68	0.008***
	3. Sociable	3.40	0.71	
	4. High cognitive ability	3.65	0.84	
Likeability				
	1. Robot expressions in response to given stimuli are understandable	3.70	0.79	
	2. Enjoyed during interactions with the robot	4.00	0.83	0.024**
	3. Want to have more interactions with the robot in the future	4.30	0.73	
Safety				
	1. Anxious	2.14	1.02	
	2. Scary	2.08	0.92	0.570
	3. Aggressive	1.85	0.99	
Robot expressions				
	1. Movement (i.e., slow, fast)	3.46	0.61	0.424
	2. Emotion-based expressions (i.e., happy, sad)	3.62	0.78	

similar to human emotions.

The proposed architecture was evaluated by a two-phase experiment that consisted of a training phase and a testing phase. Specifically, human subjects interacted with an off-the-shelf robot Pepper through a speech interface. In summary, participants enjoyed teaching the robot through interactions. Robot responses are understandable to participants. Moreover, participants showed positive feedback for the robot "cognitive ability" and other *human-likeness* characteristics. They also presented a highly positive intention for future interactions. Almost all participants perceived the lowest level of unsafe conditions and they positively agree with the movement and expressions of the robot.

REFERENCES

[1] T. Fong, I. Nourbakhsh, and K. Dautenhahn, "A survey of socially interactive robots," *Robotics and Autonomous Systems*, vol. 42, no. 3, pp. 143–166, 2003.

[2] F. Leconte, F. Ferland, and F. Michaud, "Design and integration of a spatio-temporal memory with emotional influences to categorize and recall the experiences of an autonomous mobile robot," *Autonomous Robots*, vol. 40, no. 5, pp. 831–848, 2016.

[3] Z. Kasap and N. Magnenat-Thalmann, "Building long-term relationships with virtual and robotic characters: the role of remembering," *The Visual Computer*, vol. 28, no. 1, pp. 87–97, 2012.

[4] F. Pratama, F. Mastrogiovanni, S. G. Lee, and N. Y. Chong, "Long-term knowledge acquisition using contextual information in a memory-inspired robot architecture," *Journal of Experimental & Theoretical Artificial Intelligence*, vol. 29, no. 2, pp. 313–334, 2017.

[5] J. Posner, J. A. Russell, and B. S. Peterson, "The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology," *Development and Psychopathology*, vol. 17, no. 3, pp. 715–734, 2005.

[6] A. L. Thomaz, M. Berlin, and C. Breazeal, "An embodied computational model of social referencing," in *IEEE International Workshop on Robot and Human Interactive Communication*. IEEE, 2005, pp. 591–598.

[7] T. L. Q. Dang, S. Jeong, and N. Y. Chong, "Personalized robot emotion representation through retrieval of memories," in *International Conference on Control, Automation and Robotics*. IEEE, 2017, pp. 65–70.

[8] T. L. Q. Dang, N. T. V. Tuyen, S. Jeong, and N. Y. Chong, "Encoding cultures in robot emotion representation," in *IEEE International Symposium on Robot and Human Interactive Communication*. IEEE, 2017.

[9] N. T. Viet Tuyen, S. Jeong, and N. Y. Chong, "Emotional bodily expressions for culturally competent robots through long term human-robot interaction," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Oct 2018, pp. 2008–2013.

[10] "Who is pepper?" https://www.ald.softbankrobotics.com/en/robots/pepper, 2016, accessed: 2018-08-08.

[11] B. Bruno, N. Y. Chong, H. Kamide, S. Kanoria, J. Lee, Y. Lim, A. K. Pandey, C. Papadopoulos, I. Papadopoulos, F. Pecora, A. Saffiotti, and A. Sgorbissa, "Paving the way for culturally competent robots: A position paper," in *2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, Aug 2017, pp. 553–560.

[12] A. A. Khaliq, U. Kckemann, F. Pecora, A. Saffiotti, B. Bruno, C. T. Recchiuto, A. Sgorbissa, H. Bui, and N. Y. Chong, "Culturally aware planning and execution of robot actions," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Oct 2018, pp. 326–332.

[13] R. Menicatti, C. T. Recchiuto, B. Bruno, R. Zaccaria, A. A. Khaliq, U. Kckemann, F. Pecora, A. Saffiotti, H. Bui, N. Y. Chong, Y. Lim, V. C. Pham, N. T. V. Tuyen, N. Melo, J. Lee, M. Busy, E. Lagrue, J. Montanier, A. K. Pandey, and A. Sgorbissa, "Collaborative development within a social robotic, multi-disciplinary effort: the caresses case study," in *2018 IEEE Workshop on Advanced Robotics and its Social Impacts (ARSO)*, Sep. 2018, pp. 117–124.

[14] H. Bui and N. Y. Chong, "An integrated approach to human-robot-smart environment interaction interface for ambient assisted living," in *2018 IEEE Workshop on Advanced Robotics and its Social Impacts (ARSO)*, Sep. 2018, pp. 32–37.

[15] W. Dodd and R. Gutierrez, "The role of episodic memory and emotion in a cognitive robot," in *IEEE International Workshop on Robot and Human Interactive Communication*. IEEE, 2005, pp. 692–697.

[16] R. D. Lane, L. Ryan, L. Nadel, and L. Greenberg, "Memory reconsolidation, emotional arousal, and the process of change in psychotherapy: New insights from brain science," *Behavioral and Brain Sciences*, vol. 38, 2015.

[17] V. Kuperman, Z. Estes, M. Brysbaert, and A. B. Warriner, "Emotion and language: valence and arousal affect word recognition." *Journal of Experimental Psychology: General*, vol. 143, no. 3, p. 1065, 2014.

[18] J. R. Anderson, *The adaptive character of thought*. Psychology Press, 1990.

[19] A. B. Warriner, V. Kuperman, and M. Brysbaert, "Norms of valence, arousal, and dominance for 13,915 english lemmas," *Behavior Research Methods*, vol. 45, no. 4, pp. 1191–1207, 2013.

[20] N. Haslam, S. Loughnan, Y. Kashima, and P. Bain, "Attributing and denying humanness to others," *European Review of Social Psychology*, vol. 19, no. 1, pp. 55–85, 2008.

[21] C. Bartneck, E. Croft, and D. Kulic, "Measuring the anthropomorphism, animacy, likeability, perceived intelligence and perceived safety of robots," in *Metrics for HRI workshop, technical report*, vol. 471, 2008, pp. 37–44.