Autonomous Speech Volume Control for Social Robots in a Noisy Environment Using Deep Reinforcement Learning*

Ha-Duong Bui and Nak Young Chong
School of Information Science
Japan Advanced Institute of Science and Technology
1-1 Asahidai, Nomi, Ishikawa 923-1292, Japan
{*bhduong, nakyoung*}*@jaist.ac.jp*

***Abstract*— This paper presents a novel approach to automatically adjusting the speech volume of a socially assistive humanoid robot to enhance the quality of human-robot interactions. We apply the Deep Q-learning algorithm to enable the robot to adapt to the preferences of a user in the volume of the robot's voice in social contexts. Subjective experiments were conducted to verify the validity of the proposed system. Twenty-three human subjects had social conversations with humanoid robots across various noisy environments. Participants rated their perception of the robots' voices in terms of clearness and comfortability through a questionnaire. The results show that the robot equipped with our framework outperforms other experimental robots in trials. This study confirmed the effectiveness of the proposed autonomous speech volume control system for social robots communicating with people in noisy environments.**

***Index Terms*— Social Human-Robot Interaction, Speech Volume Control, Reinforcement Learning**

I. INTRODUCTION

Nowadays, the cooperation between humans and robots is turning essential in our society, as socially assistive humanoid robots become more prevalent. Thus, the development of models that can improve the interaction between humans and robots become increasingly important. One of the main tasks in the field of Human-Robot Interaction (HRI) is to provide cognitive and affective capacities to robots by creating architectures that enable them to achieve empathic connections with users [1], [2]. In the literature, there are many studied in adapting robot behavior that was proposed to address this open-challenge. For example, in [3]–[6] the authors studied social eye gaze behavior in HRI. Several works, for instance [7]–[9], have been performed on the body and facial expression cues. Many attempts have been made [10]–[12] in order to improve the robots in terms of speech and vocal. However, an approach to adjusting the talking volume of robots in social contexts is still lacking.

*This work is 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 (AFOSRAOARD/ FA2386-19-1-4015).

Therefore, in this paper, we aim to present an innovative framework to support the humanoid robots to automatically adjust the volume of its voice based on the daily conversation contexts (the Pepper robot[1] is employed as a socially assistive robot in this research). The Deep Q-learning algorithm is adopted to solve our problem. Additionally, we conduct experimental studies to investigate the correctness of our proposed architecture.

In the following, we describe the details of our proposed autonomous speech volume control system in Section II. In Section III, we present the design of the subjective experiments, followed by the methodology used to collect and analyze the data as well as the results. Finally, we share our conclusions in Section IV.

II. PROPOSED AUTONOMOUS SPEECH VOLUME CONTROL SYSTEM

A. System Overview

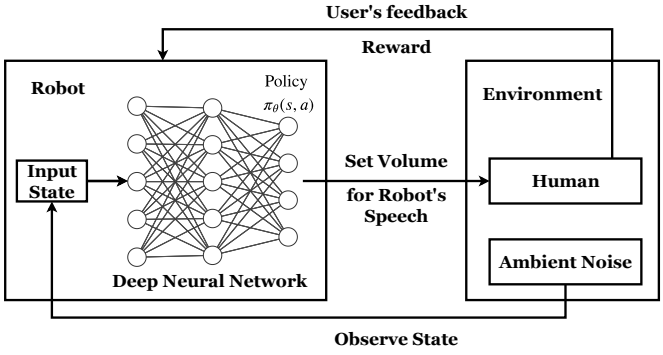


Fig. 1: System Overview

We proposed a system in order to enable a social robot to autonomously adjust its speech volume in noisy environments. Figure 1 shows the overall system architecture. Specifically, we defined our problem as a Reinforcement Learning [13] task and solved it by adapting the Deep Q-learning with Experience Replay algorithm [14]. To do so, the

[1]Pepper Robot `https://www.softbank.jp/en/robot/`

problem is formalized as the Optimal Control of a Markov Decision Process (MDP). In detail, we introduced a tuple $< S, A, P, R, \gamma >$ and a policy π, where

- S is a set of all states. A state $s \in S$ is a 4×1 vector representing the level of ambient noises and obtained by using four microphones of the robot.
- A is a set of all actions. An action $a_i \in A$ will set the volume of the robot's voice to $(i \times 5)$ percent, $i \in [0, 20]$.
- P is a transition function that returns the probability of the next state being $s\prime$ if we take action a.
- R is a reward function that returns a reward value r given as a result of taking action a in state s.

$$\text{Reward } r_t = \begin{cases} 10, & \text{if user gives positive feedbacks} \\ 0, & \text{otherwise} \end{cases}$$

- π is a policy defines the behavior of the robot. In other words, it represents the preference of user in the volume of robot's speech in social communications.

With the definition of our problem as a Markov Decision Process, we can learn the optimal policy $\pi_\theta(s,a)$ by using the Q-learning algorithm. Basically, Q-learning learns an action-value function (also known as Q function). Given a state and an action, it returns the value of taking that action. The following equation is used to update the values of Q, usually performing the update in an online manner, sampling directly from the environment each time-step,

$$Q(s,a) \leftarrow Q(s,a) + \alpha[r + \gamma \max_{a\prime} Q(s\prime, a\prime) - Q(s,a)]. \quad (1)$$

In Equation (1), the sum of the immediate reward and the discounted future reward is calculated to update the current estimate of the action-value function $Q(s,a)$. This discounted future reward is estimated by taking the maximum Q-value of all possible actions in the next state and multiplying by the discount factor γ. The Q-learning update rule gives a new estimate which is used to update the stored Q-value by taking the difference and updating proportionally to the learning rate α. [15]

With the Deep Q-learning algorithm, the neural network serves as a function approximator and parameterize the action-value function.

We employed the Deep Q-learning with Experience Replay algorithm as in Algorithm 1 to solve our defined problem.

Algorithm 1 Deep Q-learning with Experience Replay Algorithm [14]

Initialize replay memory $\mathscr{D}$ to capacity N
Initialize action-value function Q with random weights
for episode $= 1, M$ **do**
    Observe initial state s_1
    for $t = 1, T$ **do**
        With probability ε select a random action a_t
        otherwise select $a_t = \max_a Q^*(s_t, a; \theta)$
        Execute action a_t and observe reward r_t and new state s_{t+1}
        Store transition (s_t, a_t, r_t, s_{t+1}) in $\mathscr{D}$
        Sample random minibatch of transitions (s_j, a_j, r_j, s_{j+1}) from $\mathscr{D}$
        if s_{j+1} is terminal state **then**
            Set $y_j = r_j$
        else
            Set $y_j = r_j + \gamma \max_{a\prime} Q(s_{j+1}, a\prime; \theta)$
        end if
        Perform gradient descent update using $(y_j - Q(s_j, a_j; \theta))^2$ as loss
    end for
end for

B. Model Architecture

In this research, we designed the architecture of the deep neural network as in Fig. 2. The deep learning libraries, including Keras [16] and Keras-RL [17], are used to build our network. The input to our model consists of a 4×1 observed state produced by using four microphones of Pepper robot to capture the level of ambient noises. This is followed by 5 hidden layers. These hidden layers are fully-connected and consist of 100, 64, 64, 64 and 64 Rectified Linear (ReLU) units, respectively. The output layer is a fully-connected linear layer with a single output for each valid action. In our case, the number of valid actions is 21 with $action_i$ means setting the volume of robot's speakers to $(i-1) \times 5$ percent.

C. Training Details

TABLE I: The Neural Network Hyper-parameters

Parameter	Value
Capacity of replay memory N	50000
Number of training episodes M	50
Number of training steps T	1000
Random exploration probability ε	0.5
Discount factor γ	0.99
Adam learning rate	0.001

In the process of training the deep neural network, we choose the hyper-parameters in Algorithm 1 as in Table I. The network can properly optimize the action-value function $Q(s,a;\theta)$ after 5 learning episodes (approximately 5,000 learning data). In this work, the dataset was created by capturing the data of real-life conversations between the author and the robot. We recorded and processed all the information about the environment's ambient noises, the robot's actions as well as the feedback of the author. After training the neural network with 50,000 training data, we tested the network with 10,000 test cases and achieved a validation accuracy of approximately 97 percent. It proofs that the proposed system

Input Layer		Hidden Layer 1		Hidden Layer 2		Hidden Layer 3		Hidden Layer 4		Hidden Layer 5		Output Layer	
Input	Output	Input	Output	Input	Output	Input	Output	Input	Output	Input	Output	Input	Output
(None,4)	(None,4)	(None,4)	(None,100)	(None,100)	(None,64)	(None,64)	(None,64)	(None,64)	(None,64)	(None,64)	(None,64)	(None,64)	(None,21)

Fig. 2: Deep Neural Network Architecture

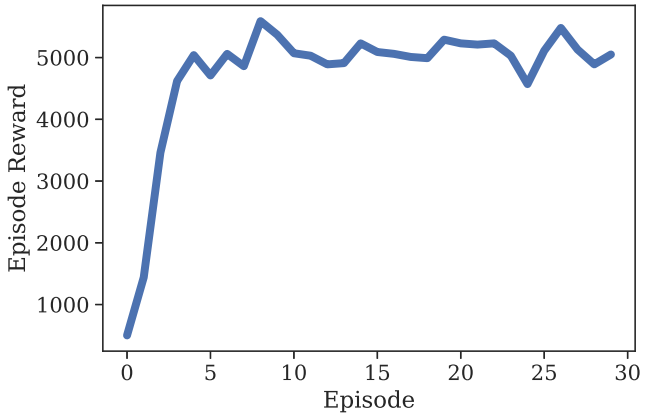


Fig. 3: Episode Reward during Training Phase

is able to learn the user's preference through daily human-robot interactions. Figure 3 shown the reward per training episode during the training phase.

III. EXPERIMENTS

In this study, subjective experiments incorporating verbal interactions between the Pepper robot and human subjects are organized to verify the validity of the proposed architecture. The details of our experimental protocol are described below with an analysis of the results obtained.

A. Design

We conducted a 4×5 experiment in which we recruited participants and let them communicate with robots of varying speech volume in environments with different reference sound levels.

Volume Level of Robot's Speakers and *Ambient Noise Level of Experimental Environments* were manipulated as two within-subjects factors.

To control the volume level of robot's speakers, we introduced four identical looking *Experimental Robots*. They are

- **Robot30** always talks in a volume of 30 %.
- **Robot50** always talks in a volume of 50 %.
- **Robot80** always talks in a volume of 80 %.
- **RobotSVA** uses the **proposed autonomous speech volume control system** to adjust its talking volume.

The sound pressure level (SPL) of the Pepper robot's speakers are described in Fig. 5a.

Manipulation of the robot's speech volume enabled testing of participants' responses to varying degrees of the robot's voice intensity. Specifically, it is served to explore any **clearness-related** differences and **comfort-related** differences in participants' evaluations.

To examine whether the proposed architecture also performs well across environments with different background noise levels, we designed five *Experimental Environments* with ambient noise levels ranging from 45 dB_{SPL} to 82 dB_{SPL} as shown in Fig. 5b. Particularly, they are

- **Environment1** is a setup of a **very quite** environment.
- **Environment2** is a setup of a **quite** environment.
- **Environment3** is a setup of a **normal** environment.
- **Environment4** is a setup of a **noisy** environment.
- **Environment5** is a setup of a **very noisy** environment.

Thus, in total, we had twenty study conditions as four *Experimental Robots* spanning five levels of ambient noises in *Experimental Environments*. Finally, this proposed protocol led us to two hypotheses H_1 and H_2,

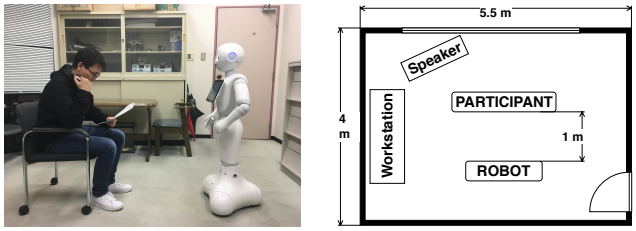
1) H_1: *Participants will be more positive toward the RobotSVA across five experimental environments measured by listening more clear to the robot's utterances.*
2) H_2: *Participants will be more positive toward the RobotSVA across five experimental environments measured by feeling more comfortable with the robot's utterances.*

B. Participants

A total of 23 participants ($n = 23$, $female = 3$, $male = 20$), with ages between 24 and 32 years old ($M = 26.52$, $SD = 2.25$), took part in the study. They were recruited among graduated students at Japan Advanced Institute of Science and Technology (JAIST) by convenience sampling. These students are from Vietnam and China and were enrolling in a Master's or Doctoral program in English. All participants confirmed that they do not have any hearing problems. We asked the participants to describe their experiences with humanoid robots with a 10-point Likert-type scale (unknown = 1; familiar = 10). It was reported that the mean number of participants' experiences with humanoid robots is 5.30 ($M = 5.30$, $SD = 3.28$). Before starting the actual tests, all participants were instructed on the experimental protocol. They could ask any questions until they can comprehend the process.

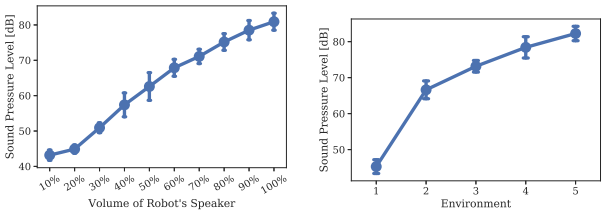
C. Settings

As illustrated in Fig. 4, the study was conducted in an ordinary room (5.5 m (W) x 4 m (D) x 2.5 m (H)) at JAIST.



(a) Side view (b) Top View

Fig. 4: Setup of Experimental Room



(a) Robot's Speakers (b) Ambient Noises in Experimental Environments

Fig. 5: Sound Pressure Levels of Robot's Speakers as well as Ambient Noises in Experimental Environments

The experimental setting comprises a humanoid robot Pepper, a 2.1 compact speaker system, and a workspace computer where all the processing takes place. We used the depth camera of the Pepper robot to calculate the distance between participants and itself. Four microphones on the robot's head were employed to listen to the surrounding sounds and provide the input to the autonomous speech volume control system. The robot interacted with participants by talking through its speakers. The sound pressure levels of the robot's speakers were calculated and reported in Fig. 5a. The compact speaker system was applied to create the different ambient noise levels in five experimental environments as in Fig. 5b by playing recorded sounds (music, people talking, white noise, etc.). The experiment was conducted in a controlled environment, including managing the temperature in the room, checking that the battery of the robots was always above 50 %, and verifying the volume of the ambient noises in experimental environments. We used the "Sound Level Meter, Class 2 NL-42"[2] for all sound pressure level measurements in this work.

D. Procedure

After welcoming a participant, each session started with a short introduction. Then, he (or she) was instructed to sit in front of the robot and was guided on how to communicate with the robot. We also requested them to only consider the loudness and to ignore other characteristics of the robot's voice (e.g. tone, pitch, range, etc.). Subsequently, the participant was asked to listen to the robot's utterances and to fill in a short survey after each trial. There were a total of 20 trials, one per study condition and each session took approximately 45 minutes.

At the beginning of each session, we shuffled the order of the appearance of the experimental robots and the experimental environments by using the Fisher-Yates algorithm. In each trial, the participant said "Start" at first. Then, the compact speaker system created the corresponding noise level in 3 seconds later. After 10 seconds, the chosen robot talked a 10-second utterance. Following a trial is a 30-second break.

A survey was given after each trial and is comprised of two questions on a 10-point Likert type scale. These questions are

1) "How clear you hear the robot's voice?" (1: cannot hear anything; 10: very clearly). With this question, we aimed to capture that if the subject can hear and acquire the transferred content from the robot (H_1).
2) "How comfortable are you with the volume of the robot's voice?" (1: annoyed; 10: very comfortable). On the other hand, the rating of how comfortable they felt about the robot's speech volume was recorded by this question (H_2).

After completing all 20 trials, the participant filled in a demographic questionnaire, and dismissed.

The standardized procedure of an experimental session is given in Algorithm 2.

E. Results

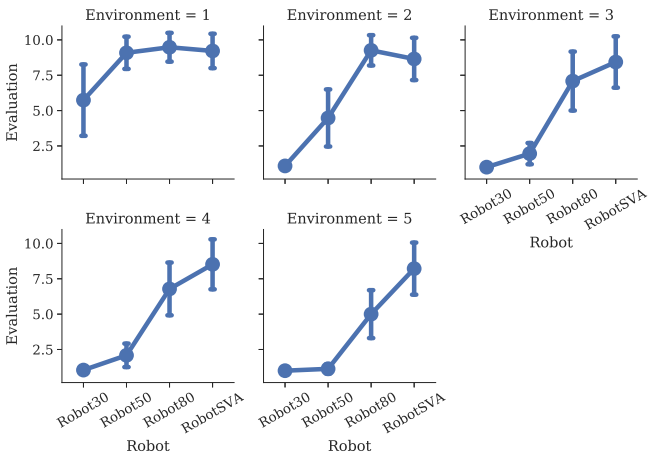


Fig. 6: Evaluations for Clearness-related through Experimental Environments (H_1)

In this research, we used the software "IBM SPSS Statistics" to analyze the data. To test our hypothesis, inferential analyses were performed at a significance level of $\alpha = 0.05$ and a Bonferroni correction for multiple comparisons was

[2]Sound Level Meter, Class 2 NL-42 https://rion-sv.com/products/10005/NL420009

TABLE II: Descriptive Statistics of Participants' Evaluations for Experimental Robots in Experimental Environments

Environment	Evaluation ($n = 23$)											
	Robot30			Robot50			Robot80			RobotSVA		
	Min	Max	Mean (SD)	Min	Max	Mean (SD)	Min	Max	Mean (SD)	Min	Max	Mean (SD)
Environment 1												
Comfort	2	8	4.74 (1.839)	4	10	7.57 (1.472)	2	9	5.43 (1.903)	6	10	8.52 (1.310)
Clearness	2	10	5.74 (2.580)	7	10	9.09 (1.164)	6	10	9.48 (1.039)	6	10	9.22 (1.242)
Environment 2												
Comfort	1	3	1.17 (0.491)	1	8	3.91 (1.703)	5	10	7.61 (1.406)	5	10	7.87 (1.604)
Clearness	1	2	1.09 (0.288)	2	9	4.48 (2.064)	7	10	9.26 (1.096)	5	10	8.65 (1.526)
Environment 3												
Comfort	1	2	1.09 (0.288)	1	3	1.65 (0.775)	2	10	6.04 (2.306)	2	10	7.83 (2.167)
Clearness	1	1	1.00 (0.000)	1	3	1.96 (0.767)	3	10	7.09 (2.130)	3	10	8.43 (1.854)
Environment 4												
Comfort	1	3	1.17 (0.491)	1	5	1.87 (1.100)	2	9	5.78 (1.833)	3	10	7.78 (1.882)
Clearness	1	2	1.04 (0.209)	1	4	2.09 (0.848)	3	10	6.78 (1.906)	4	10	8.52 (1.806)
Environment 5												
Comfort	1	2	1.04 (0.209)	1	2	1.13 (0.344)	1	8	4.39 (1.725)	3	10	7.13 (1.938)
Clearness	1	1	1.00 (0.000)	1	2	1.13 (0.344)	2	8	5.00 (1.732)	4	10	8.22 (1.882)

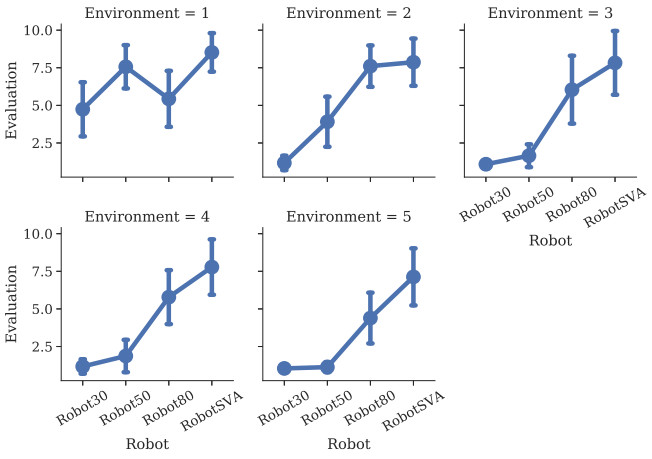


Fig. 7: Evaluations for Comfort-related through Experimental Environments (H_2)

TABLE III: Pairwise Comparisons

Measure	**(i) Robot**	**(j) Robot**	**Mean Difference (i-j)**	**Std. Error**	**Sig.[b]**
Comfort	RobotSVA	Robot30	5.983	0.296	0.000
		Robot50	4.600	0.260	0.000
		Robot80	1.974	0.185	0.000
Clearness	RobotSVA	Robot30	6.635	0.270	0.000
		Robot50	4.861	0.206	0.000
		Robot80	1.087	0.103	0.000

[b] The mean difference is significant at the 0.05 level.

applied to all post-hoc tests. Analyses were classical Two-way Repeated Measures Analysis of Variance (ANOVA). Descriptive statistics are shown in Table II, Fig. 6 and Fig. 7. Furthermore, Table III reported the post-hoc pairwise comparisons.

After performing the study, the results indicate that there are statistically significant differences between experimental

Algorithm 2 Standardized Procedure of an Experimental Session

Setup experimental room
Experimenter welcomes Participant
Experimenter gives a short introduction and guides Participant
Experimental Robots = {**Robot30**, **Robot50**, **Robot80**, **RobotSVA**}
Experimental Environments = {**Environment1**, **Environment2**, **Environment3**, **Environment4**, **Environment5**}
Shuffle *Experimental Robots*
Shuffle *Experimental Environments*
for each *Environment* in *Experimental Environments* **do**
    for each *Robot* in *Experimental Robots* **do**
        Participant says "Start"
        3 seconds pass
        Setup *Environment*
        10 seconds pass
        Robot says 10-second utterance
        Participant answers two questions in the survey
        Participant take 30-second break
    end for
end for
Participant fills in the demographic questionnaire
Experimental session is done

robots. In detail, the observations confirm that **RobotSVA** outperforms three other experimental robots through five experimental environments in terms of **clearness** and **comfort**. The participants were able to capture all the contents that **RobotSVA** talked to them with the fact that **RobotSVA** obtains excellent scores regarding **clearness** (8.22 - 9.22). Similarly, they felt very comfortable with the volume of **RobotSVA**'s voice, reflected by high scores concerning **comfort** (7.13 - 8.52). Moreover, they positively think that **RobotSVA** is able to automatically adapt its speech volume to suit with social cues in daily human-robot interactions.

Hence, we have enough evidence to accept the hypothesis H_1 and H_2. We also successfully prove the validity of our proposed autonomous speech volume control system.

IV. CONCLUSION

In this paper, we provided a new approach to automatically adjust the voice volume of a humanoid robot to improve the quality of the human-robot interactions. With subjective experiments, we are able to prove the soundness of our proposed framework. The analyzed data indicated that there are statistically significant differences between experimental robots. The observations confirmed that **RobotSVA**, the hypothetical robot supported by our proposed system, surpasses three other experimental robots through five investigational environments in terms of the **clearness** and **comfort** of the robot's voice. Specifically, the participants were able to capture all the contents that **RobotSVA** talked to them in trials. As well as, they felt very comfortable with the speech volume of the robot. **RobotSVA** successfully adapted the intensity of its voice to suit with social cues in daily conversations.

REFERENCES

[1] O. Nocentini, L. Fiorini, G. Acerbi, A. Sorrentino, G. Mancioppi, and F. Cavallo, "A survey of behavioral models for social robots," *Robotics*, vol. 8, no. 3, p. 54, 2019.

[2] E. S. Cross, R. Hortensius, and A. Wykowska, "From social brains to social robots: applying neurocognitive insights to human-robot interaction," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 374, no. 1771, p. 20180024, 2019. [Online]. Available: https://royalsocietypublishing.org/doi/abs/10.1098/rstb.2018.0024

[3] B. Mutlu, J. Forlizzi, and J. Hodgins, "A storytelling robot: Modeling and evaluation of human-like gaze behavior," in *IEEE-RAS International Conference on Humanoid Robots*, Dec 2006, pp. 518–523.

[4] C. Rich, B. Ponsler, A. Holroyd, and C. L. Sidner, "Recognizing engagement in human-robot interaction," in *ACM/IEEE International Conference on Human-Robot Interaction*, March 2010, pp. 375–382.

[5] M. Zheng, A. Moon, E. Croft, and M. Meng, "Impacts of robot head gaze on robot-to-human handovers," *International Journal of Social Robotics*, vol. 7, no. 5, pp. 783–798, 11 2015.

[6] C. Huang and B. Mutlu, "Anticipatory robot control for efficient human-robot collaboration," in *ACM/IEEE International Conference on Human-Robot Interaction*, March 2016, pp. 83–90.

[7] N. Endo, S. Momoki, M. Zecca, M. Saito, Y. Mizoguchi, K. Itoh, and A. Takanishi, "Development of whole-body emotion expression humanoid robot," in *IEEE International Conference on Robotics and Automation*, May 2008, pp. 2140–2145.

[8] Y. Kato, T. Kanda, and H. Ishiguro, "May i help you?: Design of human-like polite approaching behavior," in *ACM/IEEE International Conference on Human-Robot Interaction*, ser. HRI '15. New York, NY, USA: ACM, 2015, pp. 35–42. [Online]. Available: http://doi.acm.org/10.1145/2696454.2696463

[9] D. R. Faria, M. Vieira, F. C. C. Faria, and C. Premebida, "Affective facial expressions recognition for human-robot interaction," in *IEEE International Symposium on Robot and Human Interactive Communication*, Aug 2017, pp. 805–810.

[10] V. Chidambaram, Y.-H. Chiang, and B. Mutlu, "Designing persuasive robots: How robots might persuade people using vocal and nonverbal cues," in *ACM/IEEE International Conference on Human-Robot Interaction*, ser. HRI '12. New York, NY, USA: ACM, 2012, pp. 293–300. [Online]. Available: http://doi.acm.org/10.1145/2157689.2157798

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

[12] N. Masuyama, C. K. Loo, and M. Seera, "Personality affected robotic emotional model with associative memory for human-robot interaction," *Neurocomputing*, vol. 272, pp. 213 – 225, 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0925231217311979

[13] R. S. Sutton, A. G. Barto, *et al.*, *Reinforcement learning: An introduction*. MIT press, 1998.

[14] V. Mnih, K. Kavukcuoglu, D. Silver, A. Graves, I. Antonoglou, D. Wierstra, and M. Riedmiller, "Playing atari with deep reinforcement learning," in *NIPS Deep Learning Workshop*, 2013.

[15] J. Shannon and M. Grzes, "Reinforcement learning using augmented neural networks," *arXiv preprint arXiv:1806.07692*, 2018.

[16] F. Chollet *et al.*, "Keras," https://keras.io, 2015.

[17] M. Plappert, "keras-rl," https://github.com/keras-rl/keras-rl, 2016.