International Journal of Social Robotics (2024) 16:1523–1540
https://doi.org/10.1007/s12369-024-01142-0

Enhancing the Mobile Humanoid Robot's Emotional Expression with Affective Vertical-Oscillations

Changzeng Fu[1,2,3] · Meneses Alexis[2] · Yuichiro Yoshikawa[2] · Hiroshi Ishiguro[2]

Accepted: 1 May 2024 / Published online: 14 June 2024

Abstract
Social robots are increasingly being deployed in public environments. However, few studies have suggested ways to design robot upper body and body vertical-oscillation to enhance the robot's emotional expressions during 'walking'. This study presents a novel body-avatar interface (BoAI) enabling real-time mapping of human movements onto a robot. Using this interface, participants designed emotional upper body movements for the robot. Further, drawing inspiration from vertical oscillations in human emotional gaits, we propose emotion-specific vertical oscillation patterns for the robot. To evaluate the robot's resulting emotional expression, two video-based subjective studies were conducted with 307 and 869 participants respectively, by utilizing Mean Opinion Score (MOS) [1] and Godspeed [2] questionnaires. The results demonstrate that our proposed emotion-specific vertical oscillations significantly enhanced the robot's perceived emotional expressivity, anthropomorphism, and animacy during walking compared to neutral motion. This paper makes key contributions in designing emotional expressions for mobile robots. The BoAI and emotion-specific vertical oscillation patterns open new possibilities for improving robots' ability to express emotion during 'walking', offering possibilities to expand social robots across diverse interactive scenarios.

Keywords Robot emotional expression · Vertical oscillation · Mobile robot

1 Introduction

Emotional expression has been widely recognized as having a substantial positive impact on social interaction and interpersonal relationships for not only humans but also social robots[3–6]. According to studies in sociology, humans can perform good social norms by conveying appropriate emotions during interactions [5, 7–9]. In addition, robots can give people the impression of higher animacy, empathy, and self-disclosure [10–12] with the ability to express emotion.

With the development of robotic technology, there are many social robots that have been used in public places with locomotion functions [13–17]. To conduct better human-robot interactions (HRI), the expression of various movements, including emotional gaits, has been confirmed to be valuable for both biped robots and wheeled robots [18–22]. To enhance the affective expression of biped robots, Miura et al. [18] proposed a model of the walking pattern by imitating human motion. A similar motivation can also be seen in Izui et al.'s work [21]. They implemented human walking features to a humanoid robot (Nao) to investigate humans' success rate in recognizing the corresponding emotions from the robot's movements. Gestephe et al. [19] produced emotional gaits for the biped humanoid (Wabian-2R) to present happiness, sadness, and neutral emotion. They conducted a subjective experiment to evaluate humans' perception of the robot's expressions. Regarding wheeled robots, which offer better efficiency and stability in practical use, much research has discussed gestures and reaction timings [9, 23–27]. As for locomotion ability, Nakata et al. [28] designed a robot's emotional expression based on Laban movement analysis. Moreover, multi-modal emotional expression is also taken into account. Upper body movements and locomotion were coupled by Yagi et al. [29] to improve the android robot's

✉ Changzeng Fu
changzeng.fu@irl.sys.es.osaka-u.ac.jp;
fuchangzeng@qhd.neu.edu.cn

1 SSTC, Northeastern University, Qinhuangdao, China

2 Graduate School of Engineering Science, Osaka University, Toyonaka, Japan

3 Hebei Key Laboratory of Marine Perception Network and Data Processing, Northeastern University ai Qinhuangdao, Qinhuangdao, China

ability to portray emotion. Furthermore, Tsiourti et al. [22] discovered that a straightforward locomotion strategy can successfully express emotion when combining modalities.

However, the aforementioned studies lack clarity and diversity for robot gait-based emotional expressions. Most of them consider changes in footsteps but ignore the overall state changes (e.g. gait-induced up-down movements on the upper body), which can convey different feelings of emotion. Human body postures can indicate intentions and emotions while walking [30]. Walking patterns depend on the emotions that the person is experiencing [31–35]. For instance, a happy expression usually accompanies a sense of bouncing; an angry expression usually accompanies the feeling of stomping feet, etc [33].

As for the wheeled mobile robots that we used in this paper, it is impossible to use footsteps to generate the up-down movements due to not having feet. Therefore, inspired by [29, 36], a cylinder module is adopted to provide gait-induced vertical oscillations, simulating overall body movement. However, previous works have only implemented sinusoidal motions on the cylinder with different frequencies for different emotions. This method is limited in its ability to produce diverse overall body movement states (e.g. bouncing vs stomping) for enhanced emotional expression **(Problem 1)**. On the other hand, upper body movements are essential for expressing emotion and offer a benchmark for emotional expression with VOs. However, directly adopting human upper body motions may be inappropriate given limitations in robot degrees of freedom and appearance. Thus, studies [29, 36] designed hand-crafted upper body movements based on robots' characteristics by developers. However, basing motions solely on developers' subjective perceptions introduces individual bias. This results in movements that lack reliability and consistency as benchmarks across studies **(Problem 2)**.

To cover the mentioned limitations, the contributions of this research are:

- Different patterns of cylinder-based VO were designed to enhance the emotional expression of happiness, sadness, anger, and neutral emotion (**address problem 1**).
- A BoAI system that maps human movements to the mobile CommU in real-time was developed; participants were invited to design the mobile CommU's emotional upper body movements by using BoAI (**address problem 2**).
- The mobile CommU's emotional expression, anthropomorphism, and animacy were evaluated with published measurements (i.e. Mean Opinion Score (MOS) [1] and Godspeed [2]) video-based subjective evaluation experiments.

2 Related Works

2.1 Emotional Walking for Humans

Humans' gait patterns associated with different emotions have been widely investigated. Their findings on emotional walking patterns are concluded in Table 1. To identify emotion-specific features, Roether et al. [31] recorded the emotional gaits of different individuals, and extracted informative features from the joint-angle trajectories. The extracted features were statistically analyzed and reproduced on computer-generated characters for a subjective experiment, by which the relationship between captured features and emotions can be clarified. Lemke et al. [37] examined spatiotemporal gait parameters in patients with depression compared to healthy controls. The authors found that depressed patients walked with reduced velocity, shortened stride length, increased double limb support time, and prolonged cycle duration compared to controls. Michalak et al. [38] investigated the emotional gait patterns, including vertical movements. In their work, a musical mood induction was employed to induce sad and positive moods so that the gait patterns can be collected. The reduced walking speed, arm swing, vertical movement, and slumped posture characterize sad gait patterns, compared to happiness. Karg et al. [35] inter-individually analyzed the temporal trajectory of joints during walking with unsupervised techniques principal component analysis (PCA). From their visualization results can be inferred that happiness and anger have a larger velocity and vertical movement, while neutral and sad have a relatively small value. Kang et al. [39] collected gait data using an eight-camera optoelectronic motion capture system. They found that gait speed and smoothness increased with angry and joyful gaits compared to sadness. Specifically, vertical and anterior-posterior smoothness increased in the whole body and joints during angry and joyful walking, suggesting emotions influence movement control during gait. Halovic et al. [33] looked into more details of emotional gait patterns. They invited subjects to identify emotions by watching emotional gait videos with point-like walkers, and reported what strategies they used to conduct the judgments. They found that happiness and anger have a faster walking pace while sadness has a slower walking pace as other studies claimed. They also found that happiness is usually accompanied by a bouncy behavior and anger is always accompanied by a stomping behavior. Furthermore, for sadness, the stride length is significantly shorter than for happiness and anger, implying that the vertical movements will also be reduced. The above studies are consistent in their conclusions and can give useful insights. However, to design the upper body emotional actions of the robot based on these conclusions, some specific parameter settings are lacking (e.g. distance between arms, body/shoulder rotation, etc.). To address this problem, we developed the BoAI system and invited subjects to design the robot's emotional upper body movements.

2.2 Emotional Walking for Robots

To enhance the affective expression of robots, the emotional gait design has been considered. Destephe et al. [19] designed different set of parameters to enabled the robot to produce emotional gaits (i.e. normal, happiness, sadness). The neck pitch and waist pitch they suggested for the robot indicated that the robot faces forward for happiness while facing down for sadness. The step height and pitch range of the robot body suggested vertical movements, which are larger for happiness than sadness. Izui et al. [21] investigated emotions conveyed through gait patterns in a humanoid robot. The authors designed five emotional gaits for the Nao robot by manipulating head angle, body orientation, and walking speed. The emotional gaits were then evaluated by subjects from different cultural backgrounds. The results indicated that the emotions expressed through the robot's gaits could be recognized across cultures. Mahzoon et al. [36] examined the effects of a wheeled robot's vertical oscillation with sinusoidal function on the quality of its emotional expression. The results of the experiment indicated fast oscillations improved emotion expression with higher arousal, such as joy and anger, while slow or no oscillations were more suited to emotions with lower arousal, such as sadness. Yagi et al. [29] reproduced emotional human-like gait-induced upper body motion in a wheeled mobile android ibuki with a vertical oscillation mechanism (VOM). First, the experimenters performed and recorded the emotional gait pattern and reproduced it on the android robot. Second, they added VOM with the sinusoidal function to the robot and tested whether it could enhance the robot's emotional expression. However, both these previous studies for wheeled robots changed the amplitude and frequency of the sinusoidal function to express the corresponding emotion, without capturing the overall state (e.g. bouncing, stomping) of the body movement. In this paper, different patterns of affective cylinder-based VO are designed to enhance the corresponding emotions.

3 Methodology

3.1 Structure of the Mobile CommU

Figure 1 presents the structure of the used mobile CommU. The mobile CommU is divided into two robots: the upper robot and the bottom robot. The upper robot is a humanoid robot named CommU. This robot has 13 degrees of freedom (as shown in Fig. 1), base pitch and yaw (2 degrees of freedom), arms pitch and roll each (4 degrees of freedom) and neck pitch, roll and yaw (3 degrees of freedom), eye pitch for both eyes (1 degree of freedom), eye yaw each (2 degrees of freedom and eyelash pitch for both eyes (1 degree of freedom). The robot was designed with big eyes to show more emotional states. To receive commands for controlling CommU, an Edison board is mounted, which communicates via UART with the motors. A Websocket server, an http server, and an https server are constructed on Edison for receiving the instructions in JSON format for controlling the motor positions. The bottom robot is a mobile robot (rover) composed of 4 wheels that allow the robot to move forward and by its sides. Additionally, the robot has a cylinder that moves up and down. The CommU is placed on the top of the cylinder, allowing for vertical movements. The motors are controlled by using an ESP32 board. The Robot Operative System (ROS) was installed Nvidia Jetson Tx2 board for sending and receiving the instructions. All DoFs of the rover are controlled based on velocity control.

3.2 System for the Upper Body Movements Design

As mentioned in the previous section, the conclusions of previous works could not determine some of the parameters necessary for upper body emotional movements, such as θ_{HS}, θ_R, θ_{HR}, D_H presented in Table 2 and Fig. 2b. Instead of having developers/researchers prepare these parameters, we developed BoAI and invited subjects to participate in the design in order to prevent individual and cultural biases.

Figure 2a demonstrates the architecture of BoAI. In this system, the subject's movements will be captured frame by frame and sent into two branches for processing, one for head pose analysis and the other for body pose analysis. The analysis is realized with MediaPipe. Subsequently, we use the captured head position and the coordinate information of human joints from each frame to calculate the angle information of CommU's joints as shown in Fig. 2b. After that, the calculated angles are sent to CommU for performing the movements.

The following contents explain how we estimated and mapped head and body poses to CommU:

First, we extract the 468 landmarks from a face with MediaPipe, which returns each landmark with 3D coordinates. Then we selected 6 keypoints to estimate the head pose, which are located on the edge of the eyes (left and right canthus), tip of the nose, left and right corners of the mouth, and the jaw. Using these selected keypoints, we constructed the 2D and 3D head pose arrays to calculate the camera matrix K based on Euler's rotation theorem. We simulated the focal length of the camera matrix by simply taking the width of the image, and set the center of the image as the center of the coordinate system. As is common practice, the skew parameter of the camera matrix is set to 0. Given the 3D and 2D coordinates and the camera matrix, we can apply the Perspective-n-Point (PnP) problem to obtain the rotation matrix R as Equation 1

Table 1 The knowledge of emotional walking from previous research. The abbreviations *N, H, A, S* stand for *Neutral, Happy, Angry, Sad*, respectively

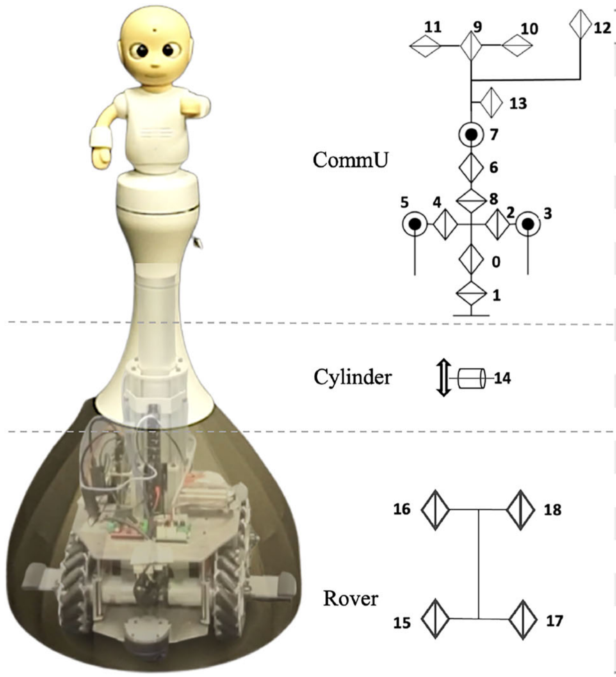
		Speed				Arm swing				Body/head orientation				Vertical movement			
Reference	Object	N	H	A	S	N	H	A	S	N	H	A	S	N	H	A	S
Roether et al. [31]	Human	–	Fast	Fast	Slow	–	Mid	Big	Small	–	Up	Up	Down	–	–	–	–
Michalak et al. [38]	Human	–	Fast	–	Slow	–	Big	–	Small	–	Mid	–	Down	–	Big	–	Small
Karg et al. [35]	Human	Mid	Fast	Fast	Slow	–	–	–	–	Up	Up	Down	Down	Mid	Big	Big	Small
Halovic et al. [33]	Human	–	Fast	Fast	Slow	–	Big	Big	Small	–	–	–	–	–	Bouncy	Stomp	Flat
Kang et al. [39]	Human	Mid	Fast	Fast	Slow	–	–	–	–	–	–	–	–	Small	Mid	Big	Small
Lemke et al. [37]	Human	–	–	–	Slow	–	–	–	Small	–	–	–	–	–	–	–	Small
Destephe et al. [19]	Robot	Mid	Fast	–	Slow	–	–	–	–	Mid	Mid	–	Down	Mid	Large	–	Small
Mahzoon et al. [36]	Robot	–	Fast	Fast	Slow	–	–	–	–	–	Mid	Side	Down	–	–	–	–
Yagi et al. [29]	Robot	–	Mid	Fast	Slow	–	–	–	–	–	Up	Lean	Down	–	Large	Mid	Small
Izui et al. [21]	Robot	Mid	Fast	Fast	Slow	–	–	–	–	Front	Up	Lean	Lean	–	–	–	–

defines, where P^c is the 2D homogeneous image point, P^w is the 3D homogeneous world point, and T is the translation matrix:

$$P^c = K\{R \mid T\}P^w = \begin{bmatrix} X_c \\ Y_c \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \quad (1)$$

Given the rotation matrix R, we conducted the RQ decomposition to obtain the head pose with orientation angles. The orientation angles are directly used to control the robot's head angle to imitate humans' head movements. θ_{HR} is included in the orientation angles.

Body pose estimation: First we extracted the positions of body joints with the help of MediaPipe, the position information is in 3D coordinates, which were used to calculate the angle parameters of CommU's joints. For θ_{AR}, we calculated the angle between the vector formed by the hand and the shoulder and the vertical direction. For θ_{AS}, we calculated the relative distance between the hand and the shoulder and used the distance from the neck to the waist as a reference to the angle parameter using the arccosine function. For θ_{HN}, we used the position of the X-axis and Z-axis of the 3D coordinates to calculate the head and neck vectors and obtained the head tilt angle. For θ_R, the position of the X-axis and Z-axis of the 3D coordinates is also used to calculate the vectors of head and shoulder to obtain the angle of body rotation. The processing speed of the BoAI system is 20 frames per second. Moreover, there are two ways of reproducing subjects' movements on CommU. One option is to let CommU imitate in real-time, so that the subject can



Joint	Name
0	Waist pitch
1	Waist yaw
2	Left hand pitch
3	Left hand roll
4	Right hand pitch
5	Right hand roll
6	Neck pitch
7	Neck roll
8	Neck yaw
9	Eyes pitch
10	Left eye yaw
11	Right eye yaw
12	Eyes lids pitch
13	Mouth pitch
14	Cylinder
15	Right front wheel
16	Right back wheel
17	Left front wheel
18	Left back wheel

Fig. 1 Hardware structure of the mobile CommU. The mobile CommU consists of an upper humanoid robot (CommU) with 13 degrees of freedom in the arms, neck, eyes, and eyelashes, which sits atop a holonomic rover allowing omnidirectional movement

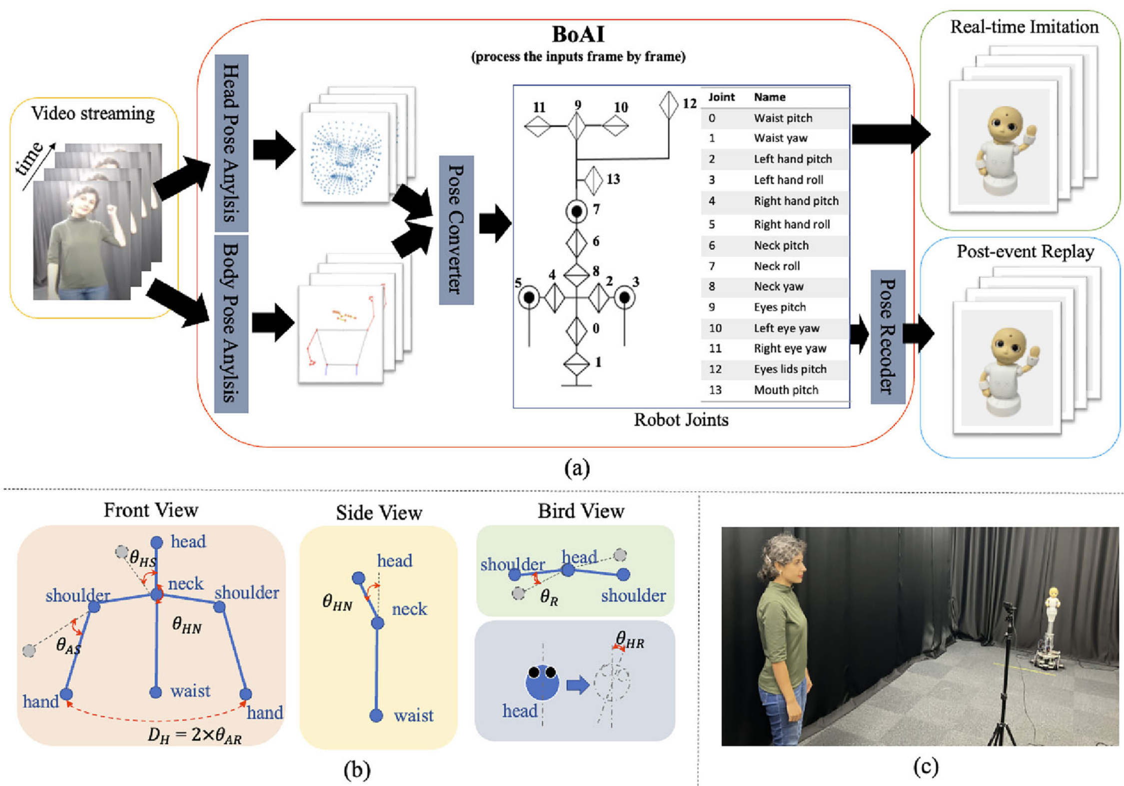


Fig. 2 **a** System architecture of the body-avatar interface (BoAI) for the upper body movement design. **b** Key parameter of upper body movements. **c** Scene of a subject deigning emotional upper body movements with BoAI

verify if CommU is behaving as expected; another option is the post-event replay function, which allows the subject to thoroughly confirm the designed emotional movements by reproducing the movement on CommU. The recorded data will be used to investigate what values of the parameters in Table 2 can be set to allow CommU to express specific emotions. The verification of the designed emotional upper body movements will be further discussed later in the paper.

3.3 Algorithm for the Affective Vertical-Oscillations

The previous works use a sinusoidal function to generate VO [29, 36], which cannot produce the overall state of bouncy and stomping movements. Therefore, as shown in Algorithm 1, we designed different VO generation methods for different emotions. Specifically, to enable the mobile CommU to exhibit a bouncy feeling, the parabolic function is considered to generate VO. Moreover, skipping behavior is a popular behavior in Japanese culture to express happiness, which consists of two parts (a big skip followed by a small skip). To implement this movement, we used two parabolic functions, one big and one small, to generate happy VO.

Algorithm 1 Algorithm for generating emotional vertical oscillations

Initialize: Stride Length $\mathcal{D}_{sl}$, Walking Velocity $\mathcal{V}_w$, Maximum Height $\mathcal{H}$, Period $\mathcal{T}$, Cylinder Velocity $\mathcal{V}_c$, Cylinder Position $\mathcal{P}$, Time Step $t \in [0, \mathcal{T}]$; $\alpha_h, \alpha_a, \alpha_s$,are weights for adjusting period/frequency.
if emotion==‘happy’ **then**
  $\mathcal{T}_1 = \mathcal{D}_{sl}/(\alpha_h \cdot \mathcal{V}_w)$; $\mathcal{T}_2 = \mathcal{T}_1/2$; $\mathcal{T} = \mathcal{T}_1 + \mathcal{T}_2$
  if $t \le \mathcal{T}_1$ **then**
    $\mathcal{P} = (-4 \cdot \mathcal{H}/\mathcal{T}_1^2) \cdot t^2 + (4 \cdot \mathcal{H}/\mathcal{T}_1) \cdot t$
    $\mathcal{V}_c = \frac{d\mathcal{P}}{dt} = (4 \cdot \mathcal{H}/\mathcal{T}_1) \cdot (-2 \cdot t/\mathcal{T}_1 + 1)$
  else
    $\mathcal{P} = (-2 \cdot \mathcal{H}/\mathcal{T}_2^2) \cdot t^2 + (2 \cdot \mathcal{H}/\mathcal{T}_2) \cdot t$
    $\mathcal{V}_c = \frac{d\mathcal{P}}{dt} = (2 \cdot \mathcal{H}/\mathcal{T}_2) \cdot (-2 \cdot t/\mathcal{T}_2 + 1)$
  end if
end if
if emotion==‘angry’ **then**
  $\mathcal{T} = \mathcal{D}_{sl}/(\alpha_a \cdot \mathcal{V}_w)$
  if t $\le \mathcal{T}/2$ **then**
    $\mathcal{P} = 2 \cdot \mathcal{H}/\mathcal{T} \cdot t$; $\mathcal{V}_c = \frac{d\mathcal{P}}{dt} = 2 \cdot \mathcal{H}/\mathcal{T}$
  else
    $\mathcal{P} = -2 \cdot \mathcal{H}/\mathcal{T} \cdot t + 2 \cdot \mathcal{H}$; $\mathcal{V}_c = \frac{d\mathcal{P}}{dt} = -2 \cdot \mathcal{H}/\mathcal{T}$
  end if
end if
if emotion==‘sad’ **then**
  $\mathcal{T} = \mathcal{D}_{sl}/(\alpha_s \cdot \mathcal{V}_w)$; $\mathcal{P} = (\sin(2\pi \cdot t/\mathcal{T}) + 1) \cdot \mathcal{H}/2$
  $\mathcal{V}_c = \frac{d\mathcal{P}}{dt} = \cos(2\pi \cdot t/\mathcal{T}) \cdot (\pi \cdot \mathcal{H}/T)$
end if

For the angry movement, the mobile CommU is supposed to present the feeling of stomping. This means that the cylinder needs to immediately change its speed direction when it is at a peak or bottom position. Therefore, we designed the movement of the cylinder on the temporal dimension as a triangular wave. For sadness, we adopted the sinusoidal function to generate VO, but reduced the amplitude and frequency to present the overall state of flatness.

For the position adjustment of the cylinder, it can be controlled with velocity by ROS. Therefore, the derivatives of the designed VO were calculated to obtain the velocity function. In general, when employing the proposed algorithm, it only needs to typically define the values of stride length $\mathcal{D}_{sl}$, walking velocity $\mathcal{V}_w$, the maximum height of cylinder $\mathcal{H}$, and the adjustment weight *alpha*. Our human generates a vertical movement at each step while walking, so $\mathcal{D}_{sl}$ is defined to model the step length of the wheeled mobile CommU, the time duration required to generate a VO can be obtained by $\mathcal{D}_{sl}/\mathcal{V}_w$, which is used as the period $\mathcal{T}$ in the algorithm. Regarding the design of *alpha*, as discussed in [40], the frequency of the vertical movement becomes larger when humans walk faster. Therefore, to present the different

Table 2 The key parameters of designing mobile CommU's upper body movements

Name	Unit	Definition
θ_{HS}	Degree	The angle between head and shoulder
θ_{HN}	Degree	The angle between head and neck
θ_R	Degree	The angle of body rotation based on yaw
θ_{HR}	Degree	The angle of head rotation based on yaw
θ_{AS}	Degree	The angle of arm swing
D_H	Degree	The distance between hands (the larger angle of the arm roll θ_{AR}, the larger distance)

velocities of emotional movements, we change the frequency to make the mobile CommU look fast or slow while keeping the horizontal movement velocity of the mobile CommU constant. This is the motivation for setting the parameter α in Algorithm 1, which is used to adjust the frequency of different emotions. The concrete parameter setting is explained in Sect. 4.

4 Experiments and Results

4.1 Experiment 1: Emotional Upper Body Movements Design

In this experiment, subjects were asked to create upper body emotional movements for CommU in accordance with its degrees of freedom. The BoAI system provides a highly open execution environment in which participants can control CommU's movements through their own bodies, allowing for a rich action space for CommU's emotional movements. Figure 2c presents a scene from this experiment. In total, 12 subjects (M = 7, F = 5, age = 26.92) were invited to participate in this experiment. The subjects came from four different countries and all had backgrounds in higher education and some basic knowledge of robotics. In the experiment, we asked subjects to design their desired emotional upper body movements for CommU. Throughout the process, they were unaware of each other's designed results, and no discussion was allowed among the subjects. The procedure and results of the experiment are described in the following section.

4.1.1 Procedure

Before starting the experiment, the experimenter explained the purpose and the procedure of the experiment to the subject, demonstrating this experiment was approved by the ethics committee in our university for research involving human subjects. The subjects' agreements to participate in the experiment were obtained. The experiment has 4 steps:

1. Firstly, subjects learn how to use BoAI to control CommU and practice it thoroughly;
2. Subsequently, subjects design upper body movements for CommU to express emotions (happy, angry, sad) with BoAI. The designed movements will be recorded for post-event replay;
3. The experimenter replays the designed movements on the mobile CommU with locomotion, and lets subjects check whether the emotion is expressed as they expected. If not, repeat step 2; if yes, save the data.
4. Finally, the experimenter conducts an interview with subjects for asking what strategy the subject used to design the mobile CommU's emotional expression movements, which is inspired by [33].

4.1.2 Results

We processed the data collected in Experiment 1. Specifically, we analyzed the means μ and standard deviation σ of each parameter presented in Table 2. Because the change of the vertical direction of the head is tiny [40], we choose only the mean value to fix θ_{HN}. For the other five parameters, we determine the valid interval by $\mu \pm \sigma$. The analyzed results are shown in Table 3. For the usage of the determined interval, we let the corresponding parameters conduct a transformation between the maximum and minimum values in each period. Thus, the mobile CommU's walking posture can be implemented (see Fig. 3). Table 4 presents the subjective suggestions for emotional upper body movements collected in the interview. The number in parentheses represents the *(number of samples / total number of samples for each item)*. These descriptions give us some insight into the movement velocity setting. For example, subjects all agreed that fast movement is needed when expressing anger, and high energy, as well as restlessness, should be presented. It suggests that a large α_a is needed to obtain a short period. For the sad emotion, subjects agreed to give a slow movement and present a low energy state, which implies the α_s should be somewhat small. Regarding happiness, there is some disagreement. Some subjects thought a fast speed was needed, while others believed a normal speed was essential to express relaxation. Based on these subjective opinions, we set α_h, α_a, α_s equal to 1.5, 2, 0.5 respectively. In order to have the same differentiation of emotional upper body movements and synchronize with VO, the same ratios were adopted. Specifically, defining a base velocity ($\mathcal{V}_b$) as a reference, then the velocities of happiness ($\mathcal{V}_b$), anger ($\mathcal{V}_a$), sadness ($\mathcal{V}_s$) are $1.5 \times \mathcal{V}_b$, $2 \times \mathcal{V}_b$, $0.5 \times \mathcal{V}_b$ respectively.

Table 3 The statistic results of subjects designed emotional upper body movements

Emotion	θ_{HS}	θ_{HN}	θ_R	θ_{HR}	θ_{AS}	$D_H/2\ (\theta_{AR})$
Happy	[−11,10]	14	[−13,13]	[−6,5]	[−52,45]	[−4, 44]
Angry	[−3,2]	12	[−9,13]	[−11,5]	[−62, 25]	[−4, 46]
Sad	[−5,13]	−10	[−12,9]	[−1,3]	[−73, −16]	[−14, 15]

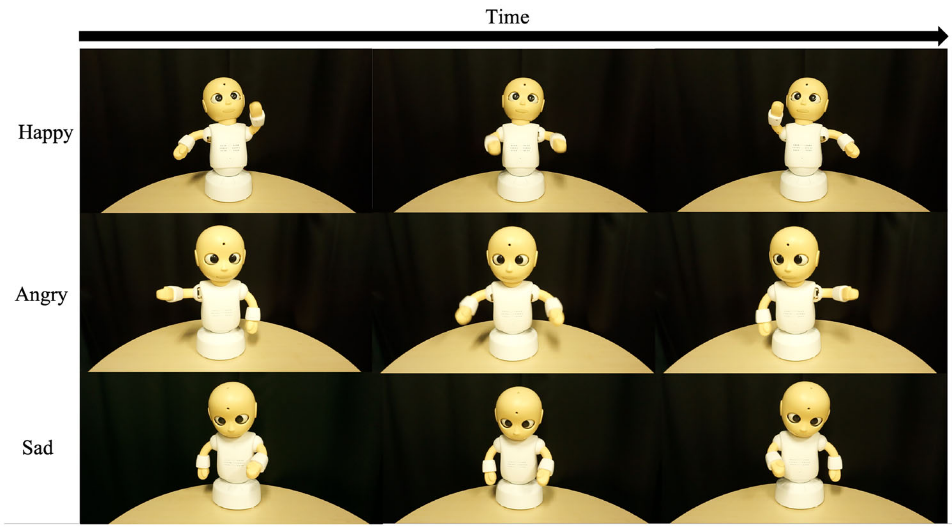


Fig. 3 The designed emotional upper body movements based on statistic results

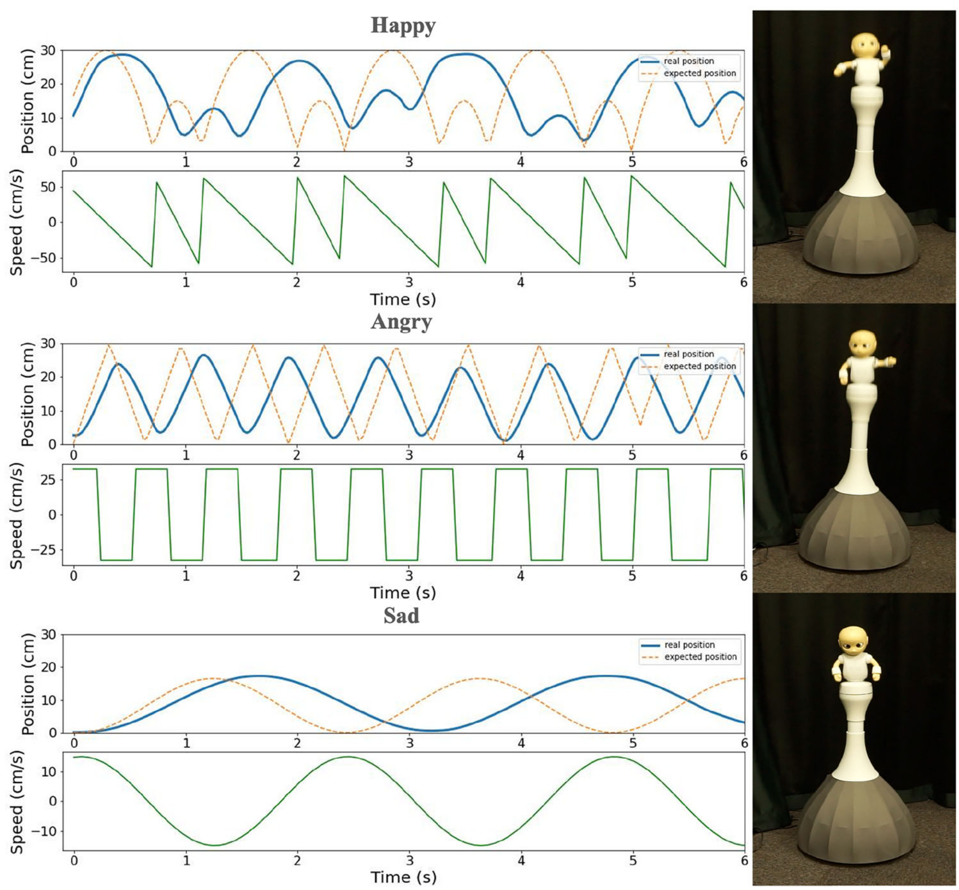


Fig. 4 Cylinder-based VO generated by the proposed algorithm. The upper row of each graph is the visualization for cylinder position, and the lower row is the one for input speeds; for the graph of positions, the dashed line indicates the expected/referred position, and the solid line indicates the real/measured position. The images on the right side are the mobile CommU equipped with the corresponding emotional movement

Table 4 Subjective strategies for emotional upper body movements collected in the interview

Emotion	Speed	Overall state
Happy	Fast(2/12); Normal(6/12); Slightly fast(4/12)	Relax(8/12); High energy(3/12); Imagine the future(1/12)
Angry	Fast(12/12)	Tense(5/21); High energy(7/21); Restless(7/21); Reluctant(2/21)
Sad	Slow(12/12)	Relax(8/12); Low energy(9/12); Calm(1/12); No interest(2/12)

4.2 Experiment 2: Verification of Designed Emotional Movements

In this experiment, we mainly want to verify the effect of VOs on the mobile robot's emotional expressions. 307 Japanese subjects in total were invited by a crowdsourcing company to evaluate the emotional expressions of the mobile CommU. Their ages ranged from 18 to 50 years old. Specifically, 104 (M = 51, F = 53, age = 28.91) participated in the evaluation of happy set, 96 (M = 47, F = 49, age = 29.88) participated in the evaluation of angry set, 107 (M = 52, F = 55, age = 29.05) participated in the evaluation of sad set. That is, for emotion sets, we used a between-subject design, and for the identical emotion expression with different VO patterns, we used a within-subject design.

4.2.1 Setup and Conditions

Given the defined values of α for each emotion in Experiment 1, the VO then can be generated. In the implementation, we set $\mathcal{D}_{sl} = 0.55\ m$, $\mathcal{V}_w = 0.2\ m/s$, $\mathcal{H} = 0.3\ m$, and $\mathcal{V}_b = 30\,deg/s$. Figure 4 shows the realized affective VO trajectories on the temporal dimension. The bottom row in each diagram represents the velocity that we input to the cylinder at each time-step. The dashed line in the top row represents the expected position of the cylinder at each moment, while the solid line represents the real position. We can see that there is a certain lag in the cylinder. But the general trajectory is in line with our expectations.

To evaluate the performance of mobile CommU's emotional expressions, we recorded the video of each condition and conducted an online experiment. The conditions set in the experiment are:

- **UBM**: This condition only uses emotional upper body movements, the VO and locomotion (forward movement) is excluded;
- **UBM+ProVO:** Based on the designed upper body movement and VO, we combined them to implement the emotional gait for the mobile CommU in the following way: (1) {UBM:happy, ProVO:happy}; (2) {UBM:angry, ProVO:angry}; (3) {UBM:sad, ProVO:sad}. The locomotion (forward movement) is implemented.
- **UBM+FlaVO:** This condition uses emotional upper body movements and flat VO, which means the forward movement is included but the vertical movement is excluded. The combinations are (1) {UBM:happy, FlaVO}; (2) {UBM:angry, FlaVO}; (3) {UBM:sad, FlaVO}.
- **UBM+SinVO:** This condition uses emotional upper body movements and locomotion (forward movement), but only implements sinusoidal VO for all upper body movements. The combinations are (1) {UBM:happy, SinVO}; (2) {UBM:angry, SinVO}; (3) {UBM:sad, SinVO}.
- **ProVO:** To investigate the impact and contribution of vertical oscillation versus upper body movement on emotional expression, we designed the ProVO condition that exclusively utilizes the proposed VO for each emotion: (1) ProVO:happy; (2) ProVO:angry; (3) ProVO:sad. Additionally, forward movement is implemented.

For each condition, we conducted a within-subject experiment, while a between-subject experiment was conducted across the emotion sets.

4.2.2 Procedure and Measurement

The experimenter first made an online questionnaire with the recorded video and prepared measurements. On the first page of the questionnaire, we gave an introduction to the experiment and informed content to ensure that each participant understood the procedures and agreed to participate. This experiment was approved by the ethics committee in our university for research involving human subjects. During the experiment, the subject should complete the following steps:

1. Firstly, subjects watch a video in its entirety (a video per condition, 5 videos per emotion set);
2. Subsequently, there are two questions for concentration check (i.e. *"describe the trajectory of mobile CommU's upper body movements"*; "describe the trajectory of the cylinder's movements");
3. After that, subjects answer the prepared questions. For emotion perception evaluation, we adopted the Mean Opinion Score (MOS). For evaluating anthropomorphism and animacy, we directly adopted the original questions from GodSpeed. The questions are shown in the following textbox.

4. Repeat steps 1 to 3 for every condition.

Regarding the measurement, it is known that MOS is a measure used in the domain of quality of experience and telecommunications engineering, representing the overall quality of a stimulus or system. It is a commonly used measure for video, audio, and audiovisual quality and emotion evaluation [11, 41–43]. Thus, we adopted the MOS with 5-level Likert scale to evaluate the mobile CommU's emotional expression quality. With the Likert scale, the higher the score, the better the quality of emotion conveyed and the more easily perceived by the subject. Moreover, to evaluate the factor of animacy, and anthropomorphism, following the previous works [12, 44], we employed the Godspeed [2]) also with 5-level Likert scale.

4.2.3 Predictions

In this experiment, we regarded the condition UBM, which only uses upper body movement, as a benchmark, and discuss the effect of VO on the mobile CommU's emotional expression. the predictions are as follows:

- **Prediction 1**: Compared to UBM, equipping the mobile CommU with the locomotion function and vertical oscillation can better enhance emotional expressions. Namely, [UBM+SinVO] and [UBM+ProVO] ≫ [UBM]; [UBM+SinVO] and [UBM+ProVO] ≫ [UBM+FlatVO].
- **Prediction 2**: Compared to UBM, equipping the mobile CommU with the proposed vertical oscillation with locomotion can better enhance emotional expressions. Namely, [ProVO] ≫ [UBM].

- ***Question for emotion evaluation with MOS:***

- From 1-5 how much do you perceive happiness from the robot?
|| neutral ① - ② - ③ - ④ - ⑤ happy ||
- From 1-5 how much do you perceive anger from the robot?
|| neutral ① - ② - ③ - ④ - ⑤ angry ||
- From 1-5 how much do you perceive sad from the robot?
|| neutral ① - ② - ③ - ④ - ⑤ sad ||

- ***Question for anthropomorphism evaluation with GodSpeed:***.

- Please rate your impression of the robot on these scales:
	fake ① - ② - ③ - ④ - ⑤ natural	
	machinelike ① - ② - ③ - ④ - ⑤ humanlike	
	unconscious ① - ② - ③ - ④ - ⑤ conscious	
	artificial ① - ② - ③ - ④ - ⑤ lifelike	
	moving rigidly ① - ② - ③ - ④ - ⑤ moving elegantly	

- ***Question for animacy evaluation with GodSpeed:***.

- Please rate your impression of the robot on these scales:
	dead ① - ② - ③ - ④ - ⑤ alive	
	stagnant ① - ② - ③ - ④ - ⑤ lively	
	mechanical ① - ② - ③ - ④ - ⑤ organic	
	artificial ① - ② - ③ - ④ - ⑤ lifelike	
	inert ① - ② - ③ - ④ - ⑤ interactive	
	apathetic ① - ② - ③ - ④ - ⑤ responsive	

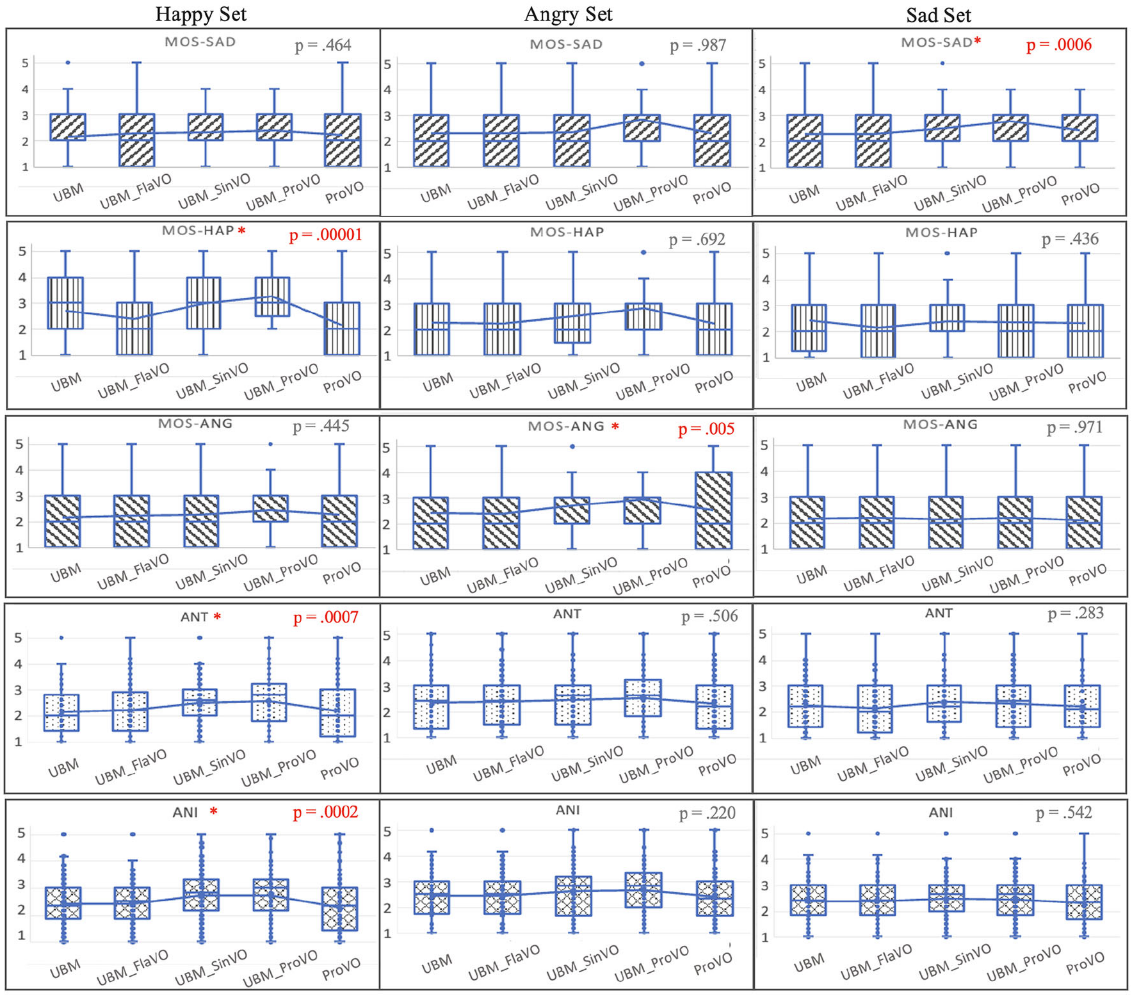


Fig. 5 Results of experiment 2. The *p* values are annotated in the upper right of the chart. Statistics with significant differences are marked in red, and those without significant differences are marked in black. ANT and ANI are the shorts for anthropomorphism and animacy. The polyline between the boxplots indicates the changing trend of the mean values among different configuration conditions

4.2.4 Results

The one-way ANOVA followed by Tukey's HSD test for multiple comparisons is adopted for the data analysis. The statistical results are shown in Fig. 5, Tables 5 and 6. Figure 5 presents boxplots illustrating the distributions of evaluation scores across conditions in Experiment 2. The *p* values from statistical significance testing are annotated in the upper right corner of each plot. Statistically significant differences between conditions are denoted in red text, while non-significant differences are in black. To further visualize the trends in the data, polylines have been overlaid connecting the mean values of each condition. In the Table 5 and Table 6, the bolded *p* value indicates $p<.017$; The bolded Cohen's d value indicates, Cohen's d>0.8, and italic Cohen's d value indicates 0.3<Cohen's d<0.8. For MOS-based emotion evaluation, because each comparison contains MOS-Happy, MOS-Angry, and MOS-Sad aspects, we penalized the *p* value with 3. The penalized *p* value with heteroscedasticity is 0.017, and we regarded a *p* value that is smaller than 0.1 but not smaller than 0.017 as showing the tendency of significant difference. For the evaluation of Anthropomorphism and Animacy, the *p* value is penalized by 2. The penalized *p* value with heteroscedasticity is 0.025, and we regarded a *p* value that is smaller than 0.1 but not smaller than 0.025 as showing the tendency of significant difference. It should be noted that sometimes the *p* value will be very small due to differences in variance between the groups but the differences in the actual means were minimal, which causes the negligible Cohen's d values. Therefore, we have not regarded this case as statistically significant.

Regarding the statistical results of the happy set, significant differences were found in happy expression ($F(4, 520) = 13.51$, $p = .00001$), anthropomorphism ($F(4,$

Table 5 Results of MOS (emotion perception evaluation) with one-way ANOVA followed with a post-hoc Tukey's HSD test. The Cohen's d results are also presented

		MOS-happy		MOS-angry		MOS-sad	
	Conditions	*p*	*d*	*p*	*d*	*p*	*d*
	UBM versus UBM+FlaVO	.233	.289	.995	.057	.908	.126
	UBM versus UBM+SinVO	.453	.227	.916	.122	.786	.174
	UBM versus UBM+ProVO	**.010**	*.482*	.361	.265	.409	.253
	UBM versus ProVO	**.010**	*.453*	.966	.089	.995	.055
Happy set	UBM+FlaVO versus UBM+SinVO	**.002**	*.521*	.991	.064	.999	.038
	UBM+FlaVO versus UBM+ProVO	**.0001**	**.816**	.606	.201	.906	.117
	UBM+FlaVO versus ProVO	.761	.169	.999	.034	.990	.060
	UBM+SinVO versus UBM+ProVO	.527	.232	.865	.135	.975	.085
	UBM+SinVO versus ProVO	**.0001**	*.683*	.999	.025	.948	.098
	UBM+ProVO versus ProVO	**.0001**	**.993**	.767	.153	.666	.172
	UBM versus UBM+FlaVO	.999	.037	.995	.051	.999	.018
	UBM versus UBM+SinVO	.995	.055	.744	.183	.999	.028
	UBM versus UBM+ProVO	.872	.131	**.016**	.470	.987	.074
	UBM versus ProVO	.998	.046	.979	.075	.000	.000
Angry set	UBM+FlaVO versus UBM+SinVO	.968	.096	.501	.239	.000	.009
	UBM+FlaVO versus UBM+ProVO	.741	.172	**.004**	*.524*	.996	.053
	UBM+FlaVO versus ProVO	.000	.009	.873	.124	.999	.017
	UBM+SinVO versus UBM+ProVO	.978	.081	.331	.327	.625	.046
	UBM+SinVO versus ProVO	.955	.104	.969	0.23	.091	.027
	UBM+ProVO versus ProVO	.703	.180	.081	*.362*	.874	.007
	UBM versus UBM+FlaVO	.386	.244	.999	.036	.999	.008
	UBM versus UBM+SinVO	.998	.042	.999	.018	.384	.267
	UBM versus UBM+ProVO	.996	.050	.998	.037	**.001**	*.522*
	UBM versus ProVO	.962	.090	.997	.045	.731	.165
Sad set	UBM+FlaVO versus UBM+SinVO	.580	.205	.995	.053	.426	.232
	UBM+FlaVO versus UBM+ProVO	.620	.194	.000	.000	**.002**	*.485*
	UBM+FlaVO versus ProVO	.804	.152	.977	.078	.771	.147
	UBM+SinVO versus UBM+ProVO	.000	.008	.995	.054	.241	*.316*
	UBM+SinVO versus ProVO	.996	.050	.92	.026	.982	.082
	UBM+ProVO versus ProVO	.998	.041	.977	.079	.071	*.367*

The bolded p value indicates the significant difference (p <.017, with penalty) with at least a small effect size; The bolded Cohen's d value indicates, Cohen's d>0.8, and italic Cohen's d value indicates 0.3<Cohen's d<0.8

520) = 4.88, $p = .0007$), and animacy ($F(4, 520) = 5.63$, $p = .0002$) aspects. No significant differences were found in angry expression ($F(4, 520) = 0.93$, $p = .445$) and sad expression ($F(4, 520) = 0.90$, $p = .464$). For the multiple comparisons of happy expression in Table 5, a significant difference can be found between UBM and UBM+ProVO ($M = 3.25$, $SD = 0.98$). It can be observed that UBM+SinVO ($M = 2.99$, $SD = 1.22$) and UBM+ProVO are significantly better than UBM+FlaVO ($M = 2.37$, $SD = 1.15$) as well as ProVO ($M = 2.18$, $SD = 1.16$) in happy expression. Moreover, a significant difference can be found between UBM and ProVO. For the multiple comparisons of anthropomorphism in Table 6, it can be seen that UBM+SinVO ($M = 2.48$, $SD = 0.85$) and UBM+ProVO ($M = 2.56$, $SD = 0.97$) can improve the anthropomorphism of the robot by comparing them to UBM ($M = 2.13$, $SD = 0.83$) with a tendency of significant difference or significant difference. Furthermore, UBM+ProVO is better than UBM+FlaVO ($M = 2.20$, $SD = 0.89$) and ProVO ($M = 2.19$, $SD = 0.98$) at presenting anthropomorphism with marginal significant difference. For the multiple comparisons of animacy in Table 6, it can

Table 6 Results of GodSpeed (anthropomorphism and animacy) with one-way ANOVA followed with a post-hoc Tukey's HSD test

		extbfAnthropomorphism		Animacy	
	Conditions	*p*	*d*	*p*	*d*
	UBM versus UBM+FlaVO	.978	.084	.000	.000
	UBM versus UBM+SinVO	.036	*.425*	.088	*.359*
	UBM versus UBM+ProVO	**.005**	.478	.056	*.379*
	UBM versus ProVO	.998	.067	.828	.141
Happy set	UBM+FlaVO versus UBM+SinVO	.153	*.327*	.087	*.366*
	UBM+FlaVO versus UBM+ProVO	.034	*.386*	.055	*.386*
	UBM+FlaVO versus ProVO	.999	.012	.832	.141
	UBM+SinVO versus UBM+ProVO	.976	.081	.999	.024
	UBM+SinVO versus ProVO	.125	*.323*	**.003**	*.464*
	UBM+ProVO versus ProVO	.026	*.381*	**.002**	*.482*
	UBM versus UBM+FlaVO	.996	.055	.000	.018
	UBM versus UBM+SinVO	.892	.128	.740	.028
	UBM versus UBM+ProVO	.671	.191	.499	.074
	UBM versus ProVO	.999	.032	.998	.000
Angry set	UBM+FlaVO versus UBM+SinVO	.983	.077	.740	.009
	UBM+FlaVO versus UBM+ProVO	.875	.141	.499	.053
	UBM+FlaVO versus ProVO	.974	.089	.998	.017
	UBM+SinVO versus UBM+ProVO	.993	.059	.995	.046
	UBM+SinVO versus ProVO	.783	.162	.545	.027
	UBM+ProVO versus ProVO	.522	.226	.315	.071
	UBM versus UBM+FlaVO	.867	.134	.000	.000
	UBM versus UBM+SinVO	.823	.141	.931	.110
	UBM versus UBM+ProVO	.974	.079	.965	.088
	UBM versus ProVO	.996	.049	.942	.105
Sad set	UBM+FlaVO versus UBM+SinVO	.257	.283	.931	.110
	UBM+FlaVO versus UBM+ProVO	.510	.214	.965	.088
	UBM+FlaVO versus ProVO	.974	.086	.942	.105
	UBM+SinVO versus UBM+ProVO	.992	.058	.999	.018
	UBM+SinVO versus ProVO	.619	.196	.526	.213

The Cohen's d results are also presented. The bolded *p* value indicates the significant difference (*p* <.025, with penalty) with at least a small effect size; The bolded Cohen's d value indicates, Cohen's d>0.8, and italic Cohen's d value indicates 0.3<Cohen's d<0.8

be observed that there is a tendency to have a significant difference between UBM+SinVO ($M = 2.71$, $SD = 0.87$) and UBM ($M = 2.41$, $SD = 0.82$), as well as UBM+SinVO and UBM+FlaVO ($M = 2.41$, $SD = 0.79$). UMB+ProVO ($M = 2.73$, $SD = 0.89$) has a tendency to be significantly better than UBM+FlaVO. Moreover, UBM+SinVO and UBM+ProVO are significantly better than ProVO ($M = 2.28$, $SD = 0.98$) in terms of presenting animacy.

Regarding the statistical results of the angry set, a significant difference was found in the angry expression ($F(4, 480) = 3.81$, $p = .005$). No significant differences were found in happy ($F(4, 480) = 0.56$, $p = .692$), sad expressions ($F(4, 480) = 0.09$, $p = .987$), anthropomorphism ($F(4, 480) = 0.83$, $p = .506$), and animacy ($F(4, 480) = 1.43$, $p = .220$). For the multiple comparisons of angry expression in Table 5, it can be seen that UBM+ProVO ($M = 2.93$, $SD = 0.95$) is significantly better than UBM ($M = 2.42$, $SD = 1.19$) and UBM+FlaVO ($M = 2.36$, $SD = 1.20$). And a tendency of significant difference can be found between UBM+ProVO and ProVO ($M = 2.51$, $SD = 1.30$).

Regarding the statistical results of the sad set, a significant difference was found in sad expression ($F(4, 535) = 4.93$, $p = .0006$). No significant differences were found in happy ($F(4, 535) = 0.95$, $p = .436$), angry expressions ($F(4, 535) = 0.13$, $p = .971$), anthropomorphism ($F(4, 535) =$

1.26, $p = .283$), and animacy ($F(4, 535) = 0.77$, $p = .542$). For the multiple comparisons of sad expression in Table 5, it can be seen that UBM+ProVO ($M = 2.79$, $SD = 0.94$) is significantly better than UBM ($M = 2.27$, $SD = 1.05$) and UBM+FlaVO ($M = 2.28$, $SD = 1.15$). And a tendency of significant difference can be found between UBM+ProVO and ProVO ($M = 2.44$, $SD = 0.98$).

4.3 Focused Experiment

The results of Experiment 2 imply the importance of VOs for the mobile robot. We would like to further explore the effect of different modes of vertical oscillations on emotion expression and verify whether different emotions require a specific vertical oscillation. Moreover, results in Experiment 2 indicate that although there is no significant difference between UBM+SinVO and UBM+ProVO in terms of emotional expression, their Cohen's d value remained at a medium size. Therefore, we believe that the comparison of these two conditions merits further validation through a focused experiment. Specifically, we used a larger sample size and redid the emotion evaluation with the MOS measurement to further investigate the effects on emotional expression between UBM+SinVO and UBM+ProVO.

In this experiment, we asked a crowdsourcing company to invite 869 Japanese individuals between the ages of 18 and 50 to assess the emotional expressions of the mobile CommU in two conditions: UBM+SinVO and UBM+ProVO. The evaluation was divided into three sets: happy, angry, and sad. Specifically, 281 subjects (M = 136, F = 145, age = 34.20) evaluated the happy set, 283 subjects (M = 133, F =1 50, age = 36.50) evaluated the angry set, and 305 subjects (M = 153, F = 152, age = 34.56) evaluated the sad set.

4.3.1 Procedures

The focused experimenter made an online questionnaire with the recorded video. On the first page of the questionnaire, we gave an introduction to the experiment and informed content to ensure that each participant understood the procedures and agreed to participate. This experiment was approved by the ethics committee in our university for research involving human subjects. During the experiment, the subject should complete three steps:

1. Firstly, subjects watch a video in its entirety (a video per condition, 5 videos per emotion set);
2. Subsequently, there are two questions for the concentration check which is identical to experiment 2;
3. After that, subjects answer the prepared questions to evaluate the robot's emotional expression. Note that anthropomorphism and animacy are not evaluated in this experiment.

4.3.2 Prediction

Prediction 3: Compared to UBM+SinVO, UBM+ProVO should obtain a higher emotional perception score for each emotion. Namely, [UBM+ProVO] $\gg$ [UBM+SinVO].

4.3.3 Results

In this experiment, a one-way ANOVA is adopted to analyze the comparisons. Because each comparison contains MOS-Happy, MOS-Angry, and MOS-Sad evaluations, we penalized the p value with 3. The penalized p value with heteroscedasticity is 0.017. We regarded a p value that is smaller than 0.1 but not smaller than 0.017 as showing the tendency of significant difference.

Figure 6 and Table 7 presents the results. For the Happy set, a significant difference can be found between UBM+SinVO ($MD = 2.67$, $SD = 1.37$) and UBM+ProVO ($MD = 3.71$, $SD = 1.62$) in the expression of happy ($F(1, 560) = 24.40$, $p = .0001$). For the Angry set, comparing to UBM+SinVO (happy:$MD = 2.24$, $SD = 1.24$)(angry:$MD = 2.00$, $SD = 0.99$), UBM+ProVO (happy:$MD = 2.01$,$SD = 0.89$)(angry:$MD = 2.47$, $SD = 1.43$) significantly decrease the expression of happy emotion ($F(1, 564) = 6.81$, $p = .009$), while significantly enhance the expression of angry emotion ($F(1, 564) = 25.79$, $p = .0001$). For the Sad set, a tendency of significant difference can be found between UBM+SinVO ($MD = 2.27$, $SD = 1.27$) and UBM+ProVO ($MD = 2.46$, $SD = 1.26$) in the expression of sad ($F(1, 610) = 4.51$, $p = .017$).

5 Discussion

5.1 Emotional Expression Design

For the design of upper body movements, we successfully obtained important parameters of joints for emotional expression through open experiments by using BoAI. Although the subjects involved in the design had different cultural backgrounds, their views on emotional expressions were consistent. This consistency across different cultures is a promising indication that the elicited emotional expressions may be broadly generalizable. However, our approach still has some limitations because the invited subjects are currently in similar life situations/environments, which perhaps limits the diversity of designed emotional movements. In future studies, we could try to recruit a more diverse pool of participants with different personalities, contextual factors, and design expertise to further expand the space of emotional expressions.

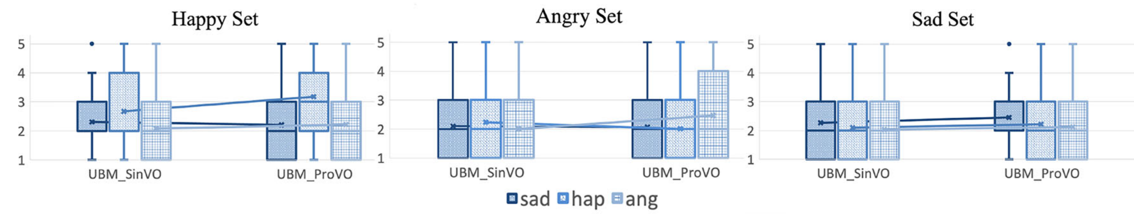


Fig. 6 Results of the focused experiment

Table 7 Results of MOS (emotion perception evaluation) with one-way ANOVA

	Conditions	extbfMOS-happy			MOS-angry			MOS-sad		
		p	*d*	*F*	*p*	*d*	*F*	*p*	*d*	*F*
Happy set	UBM+SinVO versus UBM+ProVO	**.0001**	*.59*	24.40	.108	.19	2.58	.236	.14	1.41
Angry set	UBM+SinVO versus UBM+ProVO	**.009**	*.31*	6.81	**.0001**	*.61*	25.79	.739	.03	0.11
Sad set	UBM+SinVO versus UBM+ProVO	.155	.16	2.024	.224	.14	1.48	.017	.24	4.51

The Cohen's d results are also presented. The bolded p value indicates the significant difference (p <.017, with penalty); The bolded Cohen's d value indicates, Cohen's d>0.8, and italic Cohen's d value indicates 0.3<Cohen's d<0.8

Regarding the VO design, we have successfully shaped the different emotional expressions of the robot with different movement curves by summarizing the existing findings on human emotional gait. An interesting future direction is to conduct studies to verify how well the designed VOs are recognized and interpreted during real-time human-robot interaction. This can provide insights on how to transform each expression in live interactions so that the agency and believability of the robot can be enhanced. We could evaluate factors like whether the appropriate emotion is accurately conveyed, how natural and smooth the motion appears, and how users emotionally respond to and engage with the robot.

Another limitation of the current VO design is that we relied primarily on modifying trajectory functions, frequency, and amplitude based on summaries of human emotional gaits. In future work, we could explore using data-driven methods and machine learning to automatically optimize the VO patterns based on real user feedback and interaction data. For instance, we can start with the current VO designs as a baseline and then use learning-based methods to fine-tune parameters in real-time to maximize user engagement and emotional expressiveness as evaluated through user ratings. This may allow us to achieve more natural, subtle, and fine-tuned emotional expressions through the VOs.

5.2 Subjective Emotional Expression Evaluation

Based on the results of our experiment, it appears that the VOs we proposed are effective in expressing the target emotion. We observed that significant differences in emotional expression were only found in the target emotion, while no significant differences were found in the other two emotions. The high specificity in eliciting the intended emotion provides affirmative evidence that the proposed VOs have the ability to accurately convey the intended affective states.

Compared to UBM, based on the experimental results we can claim that UBM+FlaVO failed to reinforce the presentation of emotion. It implies that only equipping the mobile CommU with a locomotion function cannot improve emotional expression. This highlights the importance of vertical oscillation as a modality for conveying emotions through gaits. Moreover, it can be seen that UBM+SinVO also failed to reinforce the presentation of emotion but UBM+ProVO successfully enhanced the emotional expression. This suggests that using only the simplistic sinusoidal function to generate mobile CommU's up-and-down movements is not sufficient to elicit stronger emotional expressions. More complex trajectory patterns seem to be needed. One potential reason that sinusoidal VO did not improve emotional expression much compared to UBM could be that the periodic up-down motion lacked the nuances and fluidity of human emotional gaits. The proposed VO may better capture these subtleties by incorporating elements like acceleration, asymmetry and irregularity seen in human gaits. These findings support Prediction 1 and provide insights into designing more effective VOs.

Regarding Prediction 2, our results show that ProVO may obtain a higher emotional perception score than UBM for the angry and sad sets but not for the happy set. A potential explanation is that for expressing happiness, the upper body movements designed through BoAI contained more salient and easily recognized emotive cues. The large motions and faster speed may have overshadowed the relative contribution of VO alone for conveying happiness. On the other hand, for subtly expressing negative emotions like anger and sadness, the more nuanced VO patterns may play a bigger role compared to neutral upper body motions. This brings up interesting questions on how the interplay between different modalities affects the expression and recognition of specific emotions.

In general, our results indicate that the combination of expressive upper body movements and specifically designed VOs works best for enhancing emotional expression. Neither UBM or ProVO alone can achieve the optimal effects. This highlights the importance of coordinated multimodal affective gait design, where different modalities likely play complementary roles. Further studies on the relative contributions and synergies between different cues for conveying particular emotions could provide useful insights for more impactful multimodal behavior generation.

To further analyze the differences between sinusoidal and proposed VOs, we conducted a focused experiment and introduced Prediction 3 that UBM+ProVO should obtain a higher emotional perception score than UBM+SinVO for each emotion. Our results partially confirm this prediction, with UBM+ProVO showing significantly higher scores than UBM+SinVO for conveying happiness and anger. These results further imply that the proposed VO patterns are better able to enhance emotional expressions compared to simple sinusoidal motion by presenting more nuances and fluidity of human emotional gaits. The higher specificity of elicited emotions also indicates that appropriate VO patterns can reduce erroneous perceptions. These findings highlight the importance of designing specific VO for different affective states.

Despite the findings that support our predictions, our proposed VO still has some limitations. We failed to enhance the robot's anthropomorphism and animacy while expressing anger and sadness. It indicates that our design of VO is not yet perfect, making the robot's movements have a certain sense of dissonance. This dissonance could arise from a lack of precise timing and coordination between the upper body motions and VO. Exploring optimal synchronization strategies could help make the overall motion more natural and harmonious. Moreover, in most of the results with significant differences, Cohen's value did not reach a large effect. This implies that our method still has room for improvement in enhancing the strength of emotional expression. More human-like VO patterns that incorporate subtle gait characteristics could potentially increase the effect size. To address this limitation, the future research direction can incorporate more nuanced human gait features into VO designs and integrate gaits with other modalities like gaze, gesture and dialogue for more holistic emotional expression.

6 Conclusion

In this paper, we discussed the design of emotional gaits with vertical oscillations (VOs) for a wheeled mobile robot, CommU. We invited participants to design the mobile CommU's emotional upper body movements via developed BoAI, and designed the affective vertical oscillations by referring to humans' emotional gaits. Video-based online experiments were conducted to evaluate the mobile CommU's emotional expressions and the presentation of anthropomorphism as well as animacy. The results suggest that simply equipping the robot with locomotion function or implementing sinusoidal VOs does not enhance emotional expressions. Designing emotion-specific VOs have the potential to enhance the mobile CommU's emotional expression. Furthermore, the robot's anthropomorphism and animacy were reinforced to some extent. Future works will further optimize affective VOs to further enhance the mobile CommU's emotional expression, anthropomorphism, and animacy.

Acknowledgements This work was supported by JSPS, Moonshot R&D Grant No. JPMJMS2011, the National Natural Science Foundation of China under Grant 62306068.

Declarations

Conflict of interest The authors declare that they have no Conflict of interest.

References

1. Katsigiannis S, Scovell J, Ramzan N, Janowski L, Corriveau P, Saad MA, Van Wallendael G (2018) Interpreting mos scores, when can users see a difference? Understanding user experience differences for photo quality. Qual User Exp 3(1):1–14
2. Bartneck C, Kulić D, Croft E, Zoghbi S (2009) Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. Int J Soc Robot 1(1):71–81
3. Bruce A, Nourbakhsh I, Simmons R (2002) The role of expressiveness and attention in human-robot interaction. In: Proceedings 2002 IEEE international conference on robotics and automation (Cat. No. 02CH37292), vol 4, pp 4138–4142. IEEE
4. Karg M, Samadani A-A, Gorbet R, Kühnlenz K, Hoey J, Kulić D (2013) Body movements for affective expression: a survey of automatic recognition and generation. IEEE Trans Affect Comput 4(4):341–359
5. Hess U, Bourgeois P (2010) You smile-i smile: emotion expression in social interaction. Biol Psychol 84(3):514–520
6. Dunn J, Brown J (1994) Affect expression in the family, children's understanding of emotions, and their interactions with others. Merrill-Palmer Quart (1982-), 120–137
7. von Scheve C (2012) The social calibration of emotion expression: an affective basis of micro-social order. Sociol Theory 30(1):1–14
8. Goffman E (2016) The presentation of self in everyday life. Social theory re-wired. Routledge, pp 482–493
9. Angelopoulos G, Rossi A, Di Napoli C, Rossi S (2022) You are in my way: non-verbal social cues for legible robot navigation

behaviors. In: 2022 IEEE/RSJ international conference on intelligent robots and systems (IROS), pp 657–662. IEEE
10. Rosenthal-von der Pütten AM, Krämer NC, Herrmann J (2018) The effects of humanlike and robot-specific affective nonverbal behavior on perception, emotion, and behavior. Int J Soc Robot 10(5):569–582
11. Fu C, Deng Q, Shen J, Mahzoon H, Ishiguro H (2022) A preliminary study on realizing human-robot mental comforting dialogue via sharing experience emotionally. Sensors 22(3):991
12. Fu C, Yoshikawa Y, Iio T, Ishiguro H (2021) Sharing experiences to help a robot present its mind and sociability. Int J Soc Robot 13(2):341–352
13. Xu S, Fang J, Hu X, Ngai E, Wang W, Guo Y, Leung VC (2022) Emotion recognition from gait analyses: Current research and future directions. IEEE Trans Comput Soc Syst 11(1):363–377
14. Hurtado M, Márquez J, Sotelo P, Cornejo J, Palomares R (2022) Mechanic design and kinematic simulation of tri-star wheeled mobile robot for covid-19 using uv-c disinfection for public transport. In: 2022 First international conference on electrical, electronics, information and communication technologies (ICEEICT), pp 1–8. IEEE
15. Iwendi C, Alqarni MA, Anajemba JH, Alfakeeh AS, Zhang Z, Bashir AK (2019) Robust navigational control of a two-wheeled self-balancing robot in a sensed environment. IEEE Access 7:82337–82348
16. Shi C, Satake S, Kanda T, Ishiguro H (2018) A robot that distributes flyers to pedestrians in a shopping mall. Int J Soc Robot 10(4):421–437
17. Mitsunaga N, Miyashita T, Ishiguro H, Kogure K, Hagita N (2006) Robovie-iv: a communication robot interacting with people daily in an office. In: 2006 IEEE/RSJ international conference on intelligent robots and systems, pp 5066–5072. IEEE
18. Miura K, Morisawa M, Kanehiro F, Kajita S, Kaneko K, Yokoi K (2011) Human-like walking with toe supporting for humanoids. In: 2011 IEEE/RSJ international conference on intelligent robots and systems, pp 4428–4435. IEEE
19. Destephe M, Brandao M, Kishi T, Zecca M, Hashimoto K, Takanishi A (2014) Emotional gait: effects on humans' perception of humanoid robots. In: The 23rd IEEE international symposium on robot and human interactive communication, pp 261–266. IEEE
20. Granados DFP, Kosuge K (2015) Design of a male-type dance partner robot for leading a physical human-robot interaction. In: 2015 IEEE international conference on mechatronics and automation (ICMA), pp 1234–1240. IEEE
21. Izui T, Milleville I, Sakka S, Venture G (2015) Expressing emotions using gait of humanoid robot. In: 2015 24th IEEE international symposium on robot and human interactive communication (RO-MAN), pp 241–245. IEEE
22. Tsiourti C, Weiss A, Wac K, Vincze M(2017) Designing emotionally expressive robots: a comparative study on the perception of communication modalities. In: Proceedings of the 5th international conference on human agent interaction, pp 213–222
23. Dautenhahn K, Werry I (2004) Towards interactive robots in autism therapy: background, motivation and challenges. Pragmat Cogn 12(1):1–35
24. Dautenhahn K (2007) Socially intelligent robots: dimensions of human-robot interaction. Philos Trans R Soc B Biol Sci 362(1480):679–704
25. Okuno Y, Kanda T, Imai M, Ishiguro H, Hagita N (2009) Providing route directions: design of robot's utterance, gesture, and timing. In: 2009 4th ACM/IEEE international conference on human-robot interaction (HRI), pp 53–60. IEEE
26. Satake S, Kanda T, Glas DF, Imai M, Ishiguro H, Hagita N(2009) How to approach humans? strategies for social robots to initiate interaction. In: Proceedings of the 4th ACM/IEEE international conference on human robot interaction, pp 109–116
27. Pandey AK, Gelin R (2018) A mass-produced sociable humanoid robot: pepper: the first machine of its kind. IEEE Robot Autom Mag 25(3):40–48
28. Nakata T, Sato T, Mori T, Mizoguchi H (1998) Expression of emotion and intention by robot body movement. In: International conference on intelligent autonomous systems 5 (IAS-5), pp 352–359
29. Yagi S, Nakata Y, Nakamura Y, Ishiguro H (2021) Perception of emotional expression of mobile humanoid robot using gait-induced upper body motion. IEEE Access 9:124793–124804
30. Aviezer H, Trope Y, Todorov A (2012) Body cues, not facial expressions, discriminate between intense positive and negative emotions. Science 338(6111):1225–1229
31. Roether CL, Omlor L, Christensen A, Giese MA (2009) Critical features for the perception of emotion from gait. J Vis 9(6):15–15
32. Gross MM, Crane EA, Fredrickson BL (2012) Effort-shape and kinematic assessment of bodily expression of emotion during gait. Hum Mov Sci 31(1):202–221
33. Halovic S, Kroos C (2018) Not all is noticed: kinematic cues of emotion-specific gait. Hum Mov Sci 57:478–488
34. Randhavane T, Bhattacharya U, Kapsaskis K, Gray K, Bera A, Manocha D (2019) Identifying emotions from walking using affective and deep features. arXiv preprint arXiv:1906.11884
35. Karg M, Kühnlenz K, Buss M (2010) Recognition of affect based on gait patterns. IEEE Trans Syst Man Cybern Part B (Cybern) 40(4):1050–1061
36. Mahzoon H, Ueda A, Yoshikawa Y, Ishiguro H (2022) Effect of robot's vertical body movement on its perceived emotion: a preliminary study on vertical oscillation and transition. Plos One 17(8):0271789
37. Lemke MR, Wendorff T, Mieth B, Buhl K, Linnemann M (2000) Spatiotemporal gait patterns during over ground locomotion in major depression compared with healthy controls. J Psychiatr Res 34(4–5):277–283
38. Michalak J, Troje NF, Fischer J, Vollmar P, Heidenreich T, Schulte D (2009) Embodiment of sadness and depression-gait patterns associated with dysphoric mood. Psychosom Med 71(5):580–587
39. Kang GE, Gross MM (2016) The effect of emotion on movement smoothness during gait in healthy young adults. J Biomech 49(16):4022–4027
40. Hirasaki A (2000) How is head and gaze stabilized while walking? (in Japanese). Biomechanism 15:107–118
41. Huynh-Thu Q, Garcia M-N, Speranza F, Corriveau P, Raake A (2010) Study of rating scales for subjective quality assessment of high-definition video. IEEE Trans Broadcast 57(1):1–14
42. ITU-T (2017) Vocabulary for performance, quality of service and quality of experience
43. Fu C, Liu C, Ishi CT, Ishiguro H (2022) An improved cyclegan-based emotional voice conversion model by augmenting temporal dependency with a transformer. Speech Commun 144:110–121
44. Fu C, Liu C, Ishi CT, Yoshikawa Y, Iio T, Ishiguro H (2021) Using an android robot to improve social connectedness by sharing recent experiences of group members in human-robot conversations. IEEE Robot Autom Lett 6(4):6670–6677

Changzeng Fu received the B.S. degree from Northeastern University, China in 2017. He received the M.S. and Ph.D. from pan in 2019 and 2022, respectively. He was with Graduate School of Engineering Science, Osaka University as an assistant professor from 2022 to 2023. Now he is currently a lecturer at Smart Sydney Technology College, Northeastern University and a guest faculty at Osaka University. His research interests include affective computing, human-robot interaction, and cognitive robots.

Alexis Meneses received a bachelor's degree in electronic engineering from San Luis Gonzaga University, Peru. He received his Master of Engineering Degree and his Ph.D. in engineering from Osaka University, Japan. He has been conducting research on affective agents which synchronize with humans. He was part of ERATO project which wanted to develop a symbiotic Human-Robot Interaction and Society 5.0 which wanted to develop a method for avoiding the limitation of the body through teleoperation. Currently, He is part of MoonShoot Project to realize an avatar-symbiotic society in which Cybernetic agents allow everyone to perform active social roles without constraint through teleoperating multiple Cybernetic Agents.

Yuichiro Yoshikawa (Member, IEEE) received the Ph.D. degree in engineering from Osaka University, Suita, Japan, in 2005. From 2005, he was a Researcher with Intelligent Robotics and Communication Laboratories, Advanced Telecommunications Research Institute International. Since 2010, he has been an Associate Professor with the Graduate School of Engineering Science, Osaka University. He is a Member of Japanese Society of Robotics, Japanese Society of Cognitive Science, the Virtual Reality Society of Japan, Japanese Society for Child and Adolescent Psychiatry, and Japanese Society of Pediatric Psychiatry and Neurology.

Hiroshi Ishiguro received a D.Eng. in systems engineering from Osaka University, Japan in 1991. He is currently Professor of Department of Systems Innovation in the Graduate School of Engineering Science at Osaka University (2009-), Distinguished Professor of Osaka University (2013-) and visiting Director (2014-) of Hiroshi Ishiguro Laboratories at the Advanced Telecommunications Research Institute and an ATR fellow. His research interests include distributed sensor systems, interactive robotics, and android science. He has published more than 300 papers in major journals and conferences, such as Robotics Research and IEEE PAMI. On the other hand, he has developed many humanoids and androids, called Robovie, Repliee, Geminoid, Telenoid, and Elfoid. These robots have been reported many times by major media, such as Discovery channel, NHK, and BBC. He has also received the best humanoid award four times in RoboCup. In 2011, he won the Osaka Cultural Award presented by the Osaka Prefectural Government and the Osaka City Government for his great contribution to the advancement of culture in Osaka. In 2015, he received the Prize for Science and Technology (Research Category) by the Minister of Education, Culture, Sports, Science and Technology (Mext). He was also awarded the Sheikh Mohammed Bin Rashid Al Maktoum Knowledge Award in Dubai in 2015. Tateisi Award in 2020.