# The Holistic Expression Factors of Emotional Motion for Non-humanoid Robots

Qisi Xie, Ding-Bang Luh*
School of Art and Design, Guangdong University of Technology, Guangzhou 510006, China

***Abstract*—The development of technology and the increasing prevalence of solitary living have transformed non-humanoid robots, such as robotic sweepers and mechanical pets, into potential sources of emotional support for individuals. Nevertheless, the majority of non-humanoid robots currently in existence are task-oriented and lack features such as facial expressions and sound. Existing research primarily emphasizes the details of human motion in robot motion design, while devoting less attention to the analysis of universal emotional expression factors and methods rooted in human recognition patterns. In our initial step, a theoretical framework and holistic expression factors were proposed based on Gestalt theory and SOR theory. These factors encompass vertical and horizontal motion direction, stimulation, and vertical repetition. Subsequently, animation simulation tests were conducted to confirm and examine the contributions of each factor to the recognition of emotional expressions. The results indicate that both vertical and horizontal movements can convey emotional valence. However, if both of them exist, there is no leading direction to the valence recognition result. When both vertical and horizontal movements are present, valence recognition is influenced by the combined effects of stimulation, vertical repetition, and movement direction. Simultaneously, non-humanoid robots can display recognizable emotional content when influenced by holistic expression factors. This framework can serve as a universal guide for emotional expression tasks in non-humanoid robots, proving the hypothesis that Gestalt theory is applicable in dynamic emotional recognition tasks. At the same time, these findings propose a new holistic perspective for designing emotional expression methods for robots.**

***Keywords*—*Human-robot interaction; robot emotion; non-humanoid robot; movement***

## I. Introduction

Due to evolving societal trends like the stay-at-home economy, the single economy, and an aging population, robots are poised to become indispensable companions for humans. These robots offer not only functional services but also emotional support to humans. Currently, non-humanoid robots are increasingly integrated into human life as the most ubiquitous non-human characters. For instance, everyday household items like robotic sweepers [1], interactive pets such as Cozmo [2] and BB-8 [3], the chicken-shaped Keepon [4] and seal-shaped PARO [5] used in hospitals for patient recovery, robot dogs that can assist in search and rescue, as well as drones used for large-scale performances [6] are all becoming increasingly prevalent. It is foreseeable that more non-humanoid robots will play a more significant and pervasive role in human life in the future [7]. However, these non-humanoid robots exhibit significant differences in appearance compared to humans, posing challenges in expressing emotions through human-like behavior imitation. Furthermore, many task-focused non-humanoid robots lack essential components like facial expressions and auditory capabilities [1]. Therefore, designing emotional motion expressions for non-humanoid robots that can effectively convey understandable emotional states to humans has become a crucial research topic.

The three primary categories of currently conducted research on emotional motion expression in robots are artificial design, artistic theory guiding, and emotional computing.

Firstly, artificial design based on robot features through the simulation of localized human body movements is the most widely used method in current research for expressing robot emotions and behaviors. For instance, robots were programmed to convey positive emotions through human-like gestures [8], such as cheering, applauding, and wave dancing for positive emotion, while negative emotions were represented using gestures like shrugging and crossing the arms. Similarly, Valenti et al. employed the Nao [9] robot to devise gestures based on human arm movements for expressing basic emotions. Johnson and Cuijpers conducted network experiments to investigate changes in the head position of robots, and found that people expect robots to lower their heads and gaze downward when expressing anger, sadness, fear, or disgust [10]. Moreover, Shimi [11] is a camera whose movements are expressed through artificially designed body postures. Keepon [4], designed with a biological appearance, enhanced emotional expression through gaze and body movements by manual design. The effectiveness of such expression relies on the resemblance between the robot and the human body, making it challenging to apply these results to the motion design of other robots.

Secondly, art theories have contributed methods for robotic expressions. The primary representative theories include animation theory [12] and dance theory. Animation theory relies on the 12 principles of animation to imbue robot motion with a realistic sense of life, such as helping drones convey intentions through path and speed during flight [13], and assisting robots in displaying behavior comprehensible to users [14]. Nevertheless, due to inherent limitations in robot flexibility, degrees of freedom, and movement speed, animation methods can hardly be fully utilized to convey robot emotional expressions. The Laban system [15] was originally used in dance research to describe the quality of motion. For example, Burton et al. [16] proposed to find emotionally similar movements from the database by using the Laban system and incorporating expressive content into a specific

*Corresponding Author

robot motion trajectory. However, when applying this system to robot research, it is necessary to identify suitable emotional expression elements tailored to each robot's unique characteristics. The selection varies based on individual characteristics and cannot be universally applied. Consequently, the application of art theories is limited to robots with specific attributes, and the absence of these detailed elements may limit the range of emotional expression capabilities.

Thirdly, emotional computing has emerged as a prevalent method in robot motion design. Leveraging advanced computer technology, machine learning algorithms extract relevant features from extensive labeled exemplar data, allowing the system to automatically generate expression trajectories. Common methods include principal component analysis (PCA) [17], factored conditional restricted Boltzmann machines (FCRBMs) [18], factored Gaussian process dynamic models (GPDM) [19], and neural network methods [20, 21], etc. Machine learning methods rely on human motion data, such as motion capture or large corpora. However, their generalization ability and scalability are limited. Even when using the same database, if the images used for training and testing are not the same or differ significantly, the final results will also be different. Therefore, machine learning approaches often exhibit relatively simplistic outcomes [22]. The majority of works still focus on single tasks such as walking, and typically involving specific structures [23], with humanoid structures being the most prevalent. Therefore, machine learning methods are limited to single objects with the same features.

In summary, current research on robot emotional expression primarily concentrates on specific or localized expressions. Consequently, these outcomes exhibit variations among individuals and lack generalizability, hindering the assurance of consistent expression results. Moreover, pertinent studies have yet to analyze the key factors influencing users' recognition of emotional expressions from the perspective of overall motion states.

Building upon the preceding discussion, this article poses a question grounded in the Gestalt effect: Is human recognition of emotional movement expression also influenced by their perception of the entire motion? Is it possible to convey emotion through the presentation of holistic expression factors?

This article aims to analyze the expression factors that affect users' emotional recognition of robots from a holistic perspective and help reduce the burden of communication on users during interactions with different non-humanoid robots. Consequently, it can facilitate rapid user acceptance of robots and foster the growth of related markets. At the same time, this research offers a novel perspective for designing emotional expression methods in robots by investigating recognition rules for emotional expression, which helps simplify the expression design of non-humanoid robots.

In our study, we initially established an experimental theoretical framework based on Gestalt theory and SOR theory, and then we derived a comprehensive set of expression factors, including vertical and horizontal motion directions, stimulation, and vertical repetitive motion. Subsequently, we explored the pivotal role played by movement direction and developed hypotheses grounded in approach-withdrawal theory and embodied emotion theory. Detailed information on this section is provided in Section II.

Then, Section III primarily focused on the effect of motion direction as the main factor on valence expression, which related to $H_1$. Section IV was dedicated to resolving and validating $H_2$ and $H_3$, which focused on analyzing the collective effects of all the holistic factors on emotional expression. Section V encompassed a comprehensive discussion of our findings, while Section VI offered a summary of the study's key outcomes.

## II. Conceptual Model and Hypotheses Development

Gestalt psychology [24] points out that people's perception towards objective objects are rooted in holistic relationships rather than specific elements. They emphasizes that the whole of anything is greater than its parts. Based this concept, we bring up the question: does human recognition of robot emotional actions also come from their perception of overall motion?

In addressing this question, we began by extracting holistic factors from Gestalt theory to analyze emotional expression. In Gestalt theory, holistic factors include similar appearance (similarity principle), potential contours (closure principle), continuation (continuity principle), proximity (proximity principle), and direction (common fate principle) et al. Most of these factors are applied to static graphics. However, the direction becomes a dynamic factor that is particularly relevant for motion. In the context of robotic movement, direction is a ubiquitous and fundamental element that remains consistent across variations in robot size, speed, and other variables. Hence, it serves as a universal overarching element for analyzing emotional expression effects. We categorized all motion directions into two main types: horizontal and vertical. In the realm of interaction, horizontal motion further subdivides into approaching and distancing from the target object, while vertical motion encompasses two distinct forms: upward for positive direction and downward for negative direction.

Then, we used SOR theory to find other holistic factors. The SOR theory suggests that stimuli trigger responses based on the internal sensations or behaviors of the organism (human), and this process involves the sequence of stimulus-individual (emotion) -trigger response [25]. Among these elements, stimulation is one of the transferable holistic factors, and response is presented through horizontal movement. By combining Gestalt theory, the expression framework employed in this study is structured as follows: stimuli (including two types of positive and negative stimuli) - vertical movement (including vertical upward, vertical downward, and repetitive movements in both directions) - response (horizontal movement approaching and moving away from the presentation decision). The relevant holistic factors include stimulation, vertical movement direction, horizontal movement direction, and vertical repetition. The theoretical framework guiding our analysis is outlined as Fig. 1.

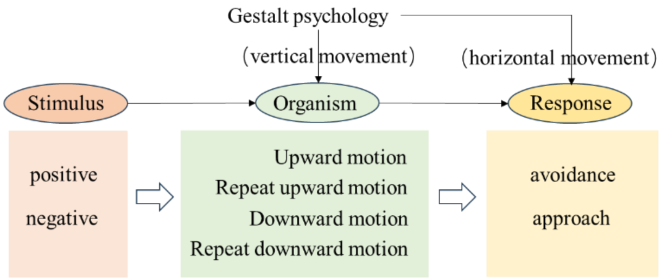


Fig. 1. Theoretical framework of the experiment.

Given the critical role of movement direction in our study, we began by analyzing the contributions of vertical and horizontal motions in the context of emotional expression to clarify their respective mechanisms for conveying emotions. According to the approach-withdrawal theory [26], we established a connection between horizontal movement and emotional valence, with approach signifying positive valence and avoidance and withdrawal signifying negative valence. The theory of embodied emotions [27] further underscores the influence of actions on the generation and perception of emotions. Consequently, positive actions can engender positive emotions, and conversely, negative actions can induce negative emotions. By extension, vertical movement aligns with the tenets of the embodied emotion theory, wherein upward motion signifies positive valence and downward motion signifies negative valence.

Notably, the valence dimension is the most basic and important aspect of the classical dimensional emotion model: Pleasure-Arousal-Dominance (PAD) model [25], which describes the range of changes in emotions from pleasant to unpleasant [28]. Compared to arousal (which represents the degree of emotional activation and is used to indicate the intensity of response to external stimuli) and dominance (which focus on the individual's control or influence over the external environment or other people, reflecting the individual's state of interaction with the environment or others), pleasant is more generalized and representative in presenting the overall emotional content. Since pleasure mainly depends on the robot's current motivation and goals [29], and it significantly informs the expresser's strategic choices in subsequent social interactions [30], the display of valence has a decisive impact on the overall emotional recognition results. Many studies on emotional expression also mainly focus on the expression of pleasure [31] [9]. In light of these considerations, we pose the question: in scenarios where both vertical and horizontal directions are concurrently present, which direction will more affect valence recognition, and how do these combined movements collectively convey emotional content?

Therefore, we have taken the following three hypotheses based on the above discussions:

$H_1$: Vertical movement has the same function as horizontal movement, which can demonstrate emotional valence.

$H_2$: When vertical and horizontal movements co-occur, one side predominates in expressing emotional valence.

$H_3$: Employing holistic factors enables non-humanoid robots to convey understandable emotional content to humans.

In this study, we first analyzed the expressive value of a single direction of motion. Subsequently, we explored the value of various holistic expression factors within the theoretical framework. These factors encompass vertical motion direction, horizontal motion direction, stimulation, and vertical repetitive motion as holistic expression factors.

## III. Factor Analysis of Valence Expression

In this study, we firstly investigated $H_1$, which posits that non-humanoid robots can convey emotional valence through vertical up-and-down movements. Additionally, we examined potential variations in recognition outcomes resulting from different types of motion modes used to achieve vertical movement.

### A. Methods

*1) Materials:* The movements that can achieve up-and-down movement in the vertical direction are translational movement and rotation. Therefore, we needed to confirm that rotating and moving on the vertical direction show the same effect. Based on the different modes of movement, we chose to use two non-humanoid robots—a spider robot and a mechanical arm—as the research objects. The spider robot allowed the body to move up and down through the support of four legs, and the robotic arm achieved the result of up and down movement by rotating.

We used the open-source animation software Blender to create simulation animations. In a total of four films, the two robots alternately went up and down to the limits of their own ranges of motion. The mechanical arm moved up and down by rotating, and the body of the spider robot moved up and down by translational motion. See Fig. 2. Then, participants evaluated the level of pleasure based on the animation contents. Data analysis used IBM SPSS Statistics 25. Materials are available at https://doi.org/10.6084/m9.figshare.23695926.v1

*2) Participants:* Participants were recruited through advertisements. 31 people (14 males, 17 females) with an age range of 18–43 (M = 28, SD = 6.5), 13 of whom had arts and humanities backgrounds, 8 had business and management backgrounds, and 9 had backgrounds in natural science and technology. Participants received a gift after the experiment. The protocol was approved by the Ethics Committee of the author's institution.

*3) Task and procedure:* Before the experiment, each participant was asked to sign a consent form and fill out demographic information. Participants were informed of task content. After the participants were ready, the simulation animation started, each animation played 3 times. After each animation finished, participants then chose the level of pleasure they felt from the robot's performance. In order to obtain more intuitive and efficient recognition results, we used the Self-Assessment Manikins (SAM) questionnaire [32] to acquire insights into how participants perceive the robot emotion [33]. Options were scored on a five-point Likert scale, with 1 being very unpleasant and 5 being very pleasant.

After the rating was completed, we started playing the next video. There were four videos in total. The experiment took an average of 7 minutes for each person.

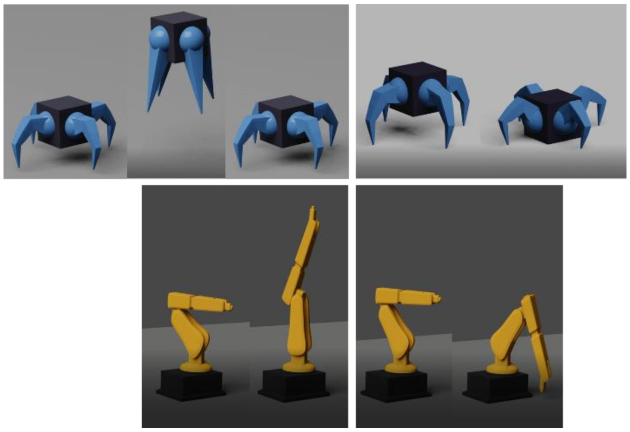

Fig. 2. Vertical movement of the robots for emotional expression. Top left: Spider robot moves upward; Top right: Spider robot moves downward; Bottom left: Mechanical arm upward movement; Bottom right: Mechanical arm downward movement.

### B. Results

To analyze the valence of the robot in the case, we plotted a mean bar chart. We used an independent sample T-test to examine the differences between different emotional groups.

Fig. 3 illustrates that there was no statistically significant difference in the recognition results for upward movement between the spider robot and the mechanical arm ($t(60) = 0.381$, $p = 0.705$). Likewise, no statistically significant difference was found in the results of downward movement between the two robots ($t(60) = 0.131$, $p = 0.896$). However, significant differences were observed in the motion recognition results of the same robot for positive and negative valence ($p < 0.001$).

To establish the range of valence recognition values within the sample's population, we performed a Z-test ($z = 2.58$) to calculate the 99% confidence interval for each parameter's estimated values, as presented in Table I.

Analysis of the data reveals that both the Spider robot and the Mechanical Arm consistently achieve valence recognition scores above 3.36 with a 99% confidence level for upward motion, while the probability of their downward motion scores falling below 2.59 was also above 99%. Emotions with scores above 3 on the 5-point scale were deemed positive, while those below 3 were considered negative. Thus, the statistical results confirmed a 99% probability of upward movement conveying positive emotions and downward movement conveying negative emotions. These findings support $H_1$, demonstrating that vertical movement serves the same purpose as horizontal movement in conveying emotional valence. Additionally, it is observed that achieving the same directional motion through various motion modes produces consistent results.

### C. Discussion

Through simulation experiments on two non-humanoid robots that achieve vertical motion through rotation and movement, we confirmed that the motion mode has no significant impact on emotion recognition results. Which means that the expression of motion in the same direction for non-human robots with different modes of motion results in the same outcome. Therefore, we can choose one of the robots as the research representative in the following direction-related research, and the results obtained can be generalized to a certain extent.

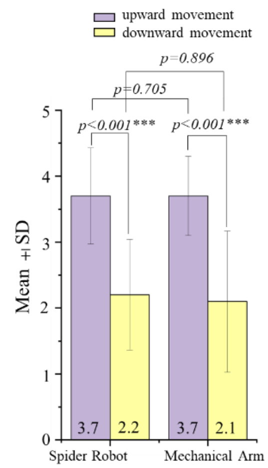


Fig. 3. Pairwise comparison between valence levels of robot emotions parameters.

TABLE I. 99% CONFIDENCE INTERVAL OF THE POPULATION TO WHICH EACH GROUP OF SAMPLES BELONGS TO THE RECOGNITION RESULTS

| **99% confidence interval** | | |
|---|---|---|
| ***movement*** | ***Spider robot*** | ***Mechanical Arm*** |
| Upward movement | [3.36, 4.04] | [3.41, 3.98] |
| Downward movement | [1.81, 2.59] | [1.61, 2.59] |

At the same time, the experiment confirms $H_1$, which suggests that vertical and horizontal movements serve the same function in demonstrating emotional valence. Accordingly, we have demonstrated the value of embodied theory in emotional action recognition, where positive behavioral actions can lead to positive emotional recognition results, and vice versa. This conclusion lays the groundwork for $H_2$ and $H_3$.

## IV. EFFECT OF HOLISTIC EXPRESSION FACTORS

It was previously verified that both vertical and horizontal movements have the function of displaying valence. Therefore, we further verified the latter two assumptions, namely $H_2$: When both vertical and horizontal movements occur, one of them dominates the display of emotional valence. $H_3$: Employing holistic factors enables non-humanoid robots to convey understandable emotional content to humans.

In order to focus on the research content, we first discussed the methods of motion combination, and then conducted experiments based on the theoretical framework and analyzed the relationship between the factors and the results.

### A. Analysis of Motion Combination Methods

Emotions often appear in complex [34] and varied forms [35]. Due to the fluid and intricate nature of emotions, not all emotional expressions exhibit distinct stages in real-life scenarios. It is also possible for concurrent of behavioral states in the 'trigger phase' and 'response phase'. Therefore, when horizontal and vertical motions occur simultaneously, there is more than one relationship between the two motions. Define the horizontal motion as $H_M$, the vertical motion as $V_M$, $t$ as the time, $n$ is the minimum interval time step between the two movements. The expressions for the two movements are as follows:

Define Action Expression as

$$\vec{A}_E(t,n) = \left(\vec{V}_M^t + \vec{H}_M^{t+n}\right) \tag{1}$$

When the two are in a sequential relationship,

$$n = 1,$$
$$\vec{A}_E(t,1) = \left(\vec{V}_M^t + \vec{H}_M^{t+1}\right) \tag{2}$$

To simplify the description, the sequential relationship described in formula (2) in the following text is denoted as $V_M \oplus H_M$.

When the two are in a parallel relationship,

$$n = 0,$$
$$\vec{A}_E(t,0) = \left(\vec{V}_M^t + \vec{H}_M^t\right) \tag{3}$$

To simplify the description, the parallel relationship described in formula (3) below is denoted as $V_M \otimes H_M$.

The two formulas describe two different relationships, with the difference being that $V_M \oplus H_M$ describes a horizontal movement that occurs after a vertical movement, and $V_M \otimes H_M$ means during the process of horizontal movement, multiple vertical movements occur simultaneously.

Since $V_M$ = {Upward movement, Downward movement}, $H_M$ = {Approach, Avoidance}, there are various types of motion combinations. The emotional valence recognition results obtained from actions with the same semantics will not change, which means when $V_M$ = Upward movement, $H_M$ = Approach, both vertical and horizontal directions exhibit positive semantics, ultimately resulting in a positive valence. Similarly, when $V_M$ = Downward movement, $H_M$ = Avoidance, the combination of two negative semantic movements results in a negative valence.

One of our research goal was to understand the role of different movement directions in emotional expression. Therefore, we mainly focused on the results obtained by combining a positive and a negative action. In addition, in order to simulate real-life scenario responses, we classified environmental stimuli into two types: positive and negative stimuli. Therefore, there were a total of eight experimental situations. See Table II.

*B. Methods*

*1) Materials:* In the first study, we confirmed that the recognition effect of both translational movement and rotation in the vertical direction was identical. Therefore, we only used one of the robots in this experiment. We chose the spider robot for analysis since it was equipped to move both vertically and horizontally at the same time. Similarly, we also used the open-source software Blender to model animations. Participants filled out questionnaires after watching the animations. In order to minimize the influence of environmental colors on the recognition results, we set all backgrounds to a neutral gray shade.

TABLE II. SUMMARY OF EXPERIMENTAL GROUP

| Stimulation | Motion relationship | Contents | Abbreviation |
|---|---|---|---|
| Positive | $V_M \oplus H_M$ | Upward movement ⊕ Avoidance | $P:U_M \oplus A_v$ |
| | | Downward movement ⊕ Approach | $P:D_M \oplus A_p$ |
| | $V_M \otimes H_M$ | Upward movement ⊗ Avoidance | $P:U_M \otimes A_v$ |
| | | Downward movement ⊗ Approach | $P:D_M \otimes A_p$ |
| Negative | $V_M \oplus H_M$ | Upward movement ⊕ Avoidance | $N:U_M \oplus A_v$ |
| | | Downward movement ⊕ Approach | $N:D_M \oplus A_p$ |
| | $V_M \otimes H_M$ | Upward movement ⊗ Avoidance | $N:U_M \otimes A_v$ |
| | | Downward movement ⊗ Approach | $N:D_M \otimes A_p$ |

The focus of this study was solely on exploring the impact of motion direction. To minimize interference, all videos maintained consistent speed, motion distance, and fixed motion amplitude (the highest reaching position of the spider robot body was 50.7cm, the middle height position was 30.6cm, and the lowest height position was 15.0cm). The distinction arises from the fact that movements that occur in parallel will experience multiple vertical up and down movements, resulting in variations in the total duration. A total of 8 videos were included, see Fig. 4. To mitigate mutual influence between similar movement expressions, the videos were played back in the following order: 1) $N:U_M \oplus A_v$, 2) $N:D_M \oplus A_p$, 3) $N:U_M \otimes A_v$, 4) $N:D_M \otimes A_p$. Then, display videos in the same order under positive stimuli. Materials are available at https://doi.org/10.6084/m9.figshare.23695926.v1

*2) Participants:* Participants were recruited through advertisements. 35 people (20 males, 15 females) from school. The age range was 18-42 (M = 27, SD = 5.9), 11 were from arts and humanities background, 6 were from business background, and 18 were from science and technology background. Participants received a gift after the experiment. The protocol was approved by the Ethics Committee of the author's institution.

*3) Procedure:* Before the experiment, each participant was asked to sign a consent form and fill out demographic information. Participants were informed of task content. After the participants were ready, the experiment began.

Firstly, an introduction of the video was showed: "The following is the emotional expression of the robot after being praised (positive stimulus) /criticized (negative stimulus) by the owner. Please choose according to the requirements." Secondly, the animations were played. The animation began with the host's expression, with a smile representing positive stimulation and an angry expression representing negative stimulation. Subsequently, the robot actions were presented. See Fig. 4. Each animation played 3 times.

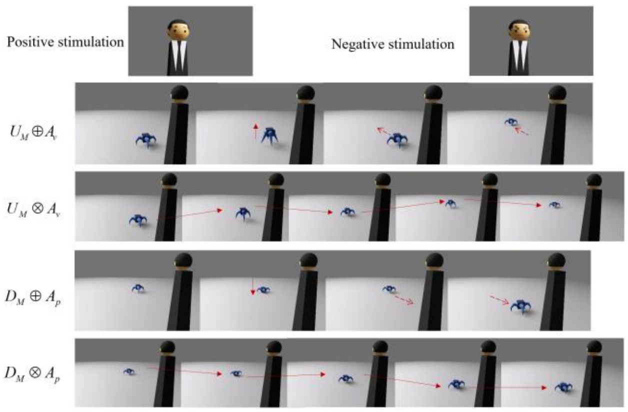


Fig. 4. Screenshots of the animations used in the experiment.

After viewing each video, participants were required to choose the pleasure level of the robot's performance as their initial response. We also used the SAM questionnaire [32]. The options were rated on a five-point Likert scale, ranging from one (very unpleasant) to five (very pleasant). Secondly, the content of different PAD emotional spaces was described using Gebhard's classification method [36], which associates various types of emotions with each emotional space. Subsequently, participants selected the emotional types of robot actions observed in the video based on their own cognition. The specific emotional options correspond to the emotional space as shown in Table III. Participants were only able to view the available options and their corresponding emotional types without direct visibility of the associated PAD emotional space. After the questionnaire ended, we continued to play the next video. The experiment took an average of 20 minutes for each person.

## C. Results

*1) Pleasure:* To comprehend the roles played by different factors in the expression process, we employed Multi-way ANOVA to examine participants' recognition of the pleasure conveyed by non-humanoid robots. Independent variables included stimuli, motion groups (Group 1: Upward movement and Avoidance, Group 2: Downward movement and Approach), and vertical repetition, while the dependent variable was recognition pleasure. Table IV displays participants' chosen pleasure levels for each group.

The results indicate that, firstly, different stimuli had a significant impact on pleasure recognition, and the main effect of the stimuli was statistically significant ($F = 40.381$，$df = 1$，$P < .001$，$\eta^2 = .129$). Secondly, different combinations of motion have no significant difference in recognition results ($F = .248$，$df = 1$，$P = .619$). Thirdly, vertical repetition significantly influenced valence recognition, and the main effect of repetition was statistically significant ($F = 10.495$，$df = 1$，$P = .001$，$\eta^2 = .037$). Finally, there was a significant interaction effect between the motion group and vertical repetition ($F = 5.030$，$df = 1$，$P = .026$，$\eta^2 = .018$). These findings were reaffirmed through independent sample T-tests for inter-group comparisons, as shown in the bar chart in Fig. 5.

TABLE III. EMOTIONAL OPTIONS AND EMOTIONAL SPACE CORRESPONDENCE TABLE

| Emotional choices | PAD space | Emotional choices | PAD space |
|---|---|---|---|
| A gratitude, liking | +P+A-D | B docile | +P-A-D |
| C pride, HappyFor | +P+A+D | D relief, relaxed | +P-A+D |
| E anger，hate | -P+A+D | F disdainful, reproach | -P-A+D |
| G shame, fear | -P+A-D | H pity, bored | -P-A-D |
| I Others （fill in the content） | | | |

TABLE IV. STATISTICS OF THE NUMBER OF PEOPLE SELECTED FOR EACH GROUP'S PLEASURE LEVEL

| | Very un-pleasure | Un-pleasure | Neutral | Pleasure | Very pleasure |
|---|---|---|---|---|---|
| $P:U_M \oplus A_v$ | 1 (2.9%) | 5 (14.3%) | 13 (37.1%) | 14 (40%) | 2 (5.7%) |
| $P:U_M \otimes A_v$ | 2 (6%) | 3 (9%) | 4 (11%) | 11 (31%) | 15 (43%) |
| $N:U_M \oplus A_v$ | 3 (9%) | 17 (48%) | 10 (29%) | 4 (11%) | 1 (3%) |
| $N:U_M \otimes A_v$ | 3 (9%) | 11(31.5%) | 5 (14%) | 11 (31.5%) | 5 (14%) |
| $P:D_M \oplus A_p$ | 0 (0%) | 5 (14.2%) | 13 (37.1%) | 13 (37.1%) | 4 (11.6%) |
| $P:D_M \otimes A_p$ | 1 (2.9%) | 3 (8.5%) | 8 (22.9%) | 23 (65.7%) | 0 (0%) |
| $N:D_M \oplus A_p$ | 1 (2.9%) | 11 (31.5%) | 18 (51.4%) | 5 (14.2%) | 0 (0%) |
| $N:D_M \otimes A_p$ | 2 (6%) | 8 (23%) | 15 (43%) | 10 (28%) | 0 (0%) |

These results address $H_2$, which showed that the presence of both vertical and horizontal movements does not result in a dominant emotional valence expression. Recognition results are collectively influenced by stimuli, vertical repetition, and the interaction between various movement directions and vertical repetition. Thus, $H_2$ is unsupported.

To further assess the interactive effects of the exercise group and vertical repetition, we conducted a simple effect analysis. The analysis revealed a significant simple effect of vertical repetition in the upward movement and avoidance group ($F = 15.03$, $df = 1$, $P < .001$), but not in the downward movement and approach group ($F = 0.497$, $df = 1$, $P = .482$). Subsequently, we conducted independent sample t-tests between groups once more, resulting in consistent results. See Fig. 6.

In the upward movement and avoidance group, both under positive stimulation ($t(68) = -2.59$, $P = 0.012$) and negative stimulation ($t(68) = -2.28$, $P = 0.026$), vertical repetition significantly influenced the results. There was no significant difference in the results between vertical repetitions group and the downward movement and approach group under positive stimulation (t (68) = -0.286, P = 0.776) and negative stimulation (t (68) = -0.891, P = 0.376).

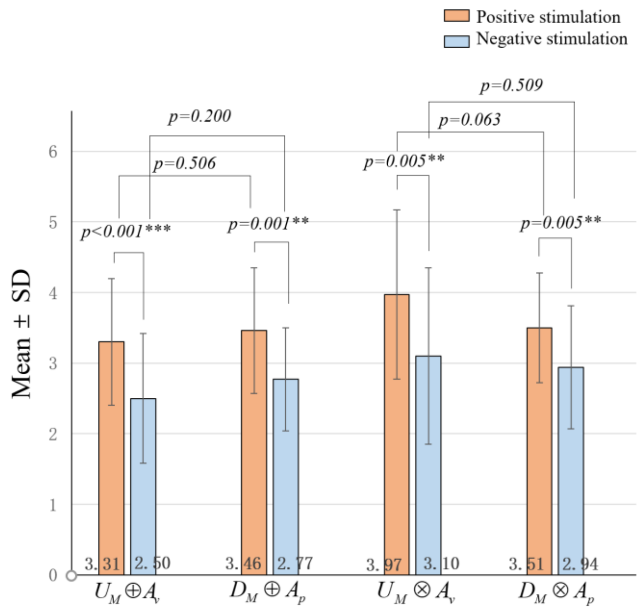


Fig. 5. Pairwise comparison of the average value of pleasure recognition under positive and negative stimuli in four groups.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

2) *Analysis of specific emotional contents:* To address $H_3$, we analyzed the specific emotional types chosen by participants for each video. The specific content corresponding to each option is shown in Table II, and the results of the selected number of people are shown in Table V.

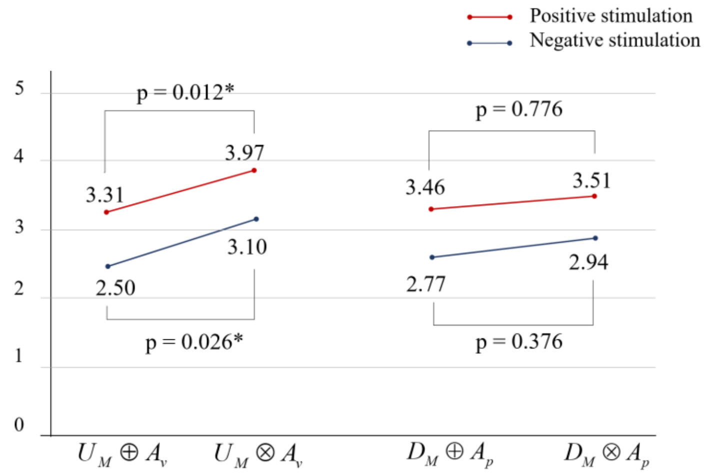


Fig. 6. Comparison of mean values between single and repeated movements in the vertical direction.

The chi-squared goodness-of-fit test results indicated that the participants' choice differed significantly from a uniform distribution. The answers with the highest number of votes were generally statistically significant. The specific results of emotion recognition are shown in Table VI.

According to the findings, non-humanoid robots' emotions can be expressed by using holistic expression factors, and answers can be acquired that are both concise and relatively unified. Thus, $H_3$ is supported.

TABLE V. NUMBER OF PEOPLE SELECTED FOR EACH EMOTIONAL CHOICE

| | $\mathbf{P:U_M \oplus A_v}$ | $\mathbf{P:U_M \otimes A_v}$ | $\mathbf{N:U_M \oplus A_v}$ | $\mathbf{N:U_M \otimes A_v}$ | $\mathbf{P:D_M \oplus A_p}$ | $\mathbf{P:D_M \otimes A_p}$ | $\mathbf{N:D_M \oplus A_p}$ | $\mathbf{N:D_M \otimes A_p}$ |
|---|---|---|---|---|---|---|---|---|
| **A (+P+A-D)** | 5 | 6 | 1 | 1 | 8 | 3 | 1 | 2 |
| **B (+P-A-D)** | 1 | 2 | 4 | 1 | 8 | 11 | 19 | 11 |
| **C (+P+A+D)** | 4 | 13 | 1 | 4 | 4 | 2 | 0 | 3 |
| **D (+P-A+D)** | 12 | 9 | 2 | 11 | 10 | 13 | 3 | 1 |
| **E (-P+A+D)** | 4 | 5 | 5 | 9 | 0 | 0 | 0 | 1 |
| **F (-P-A+D)** | 0 | 0 | 1 | 1 | 1 | 0 | 1 | 13 |
| **G (-P+A-D)** | 7 | 0 | 20 | 7 | 4 | 5 | 11 | 0 |
| **H (-P-A-D)** | 2 | 0 | 1 | 1 | 0 | 1 | 0 | 1 |
| **I Others** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 no sure |
| **Sig.** | $X^2 = 5$, p = .014 | $X^2 = 4$, p = .040 | $X^2 = 4$, p = .000 | $X^2 = 4$, p = .246 | $X^2 = 3$, p = .004 | $X^2 = 5$, p = .001 | $X^2 = 3$, p = .000 | $X^2 = 4$, p = .004 |

TABLE VI. SIGNIFICANT EMOTIONAL OPTIONS IN EACH GROUP'S RECOGNITION RESULTS

| Group | PAD result | Emotion content |
|---|---|---|
| $P:U_M \oplus A_v$ | D (+P-A+D) | relief, relaxed |
| $P:U_M \otimes A_v$ | C (+P+A+D) | pride, Happy For |
| $N:U_M \oplus A_v$ | G (-P+A-D) | shame, fear |
| $N:U_M \otimes A_v$ | -- | -- |
| $P:D_M \oplus A_p$ | D (+P-A+D) | relief, relaxed |
| $P:D_M \otimes A_p$ | D (+P-A+D) | relief, relaxed |
| $N:D_M \oplus A_p$ | B (+P-A-D) | docile |
| $N:D_M \otimes A_p$ | F (-P-A+D) | Disdainful, reproach |

### D. Discussion

This experiment firstly invalidated $H_2$, indicating no dominant direction between vertical and horizontal movements in emotional valence recognition. In order to understand the value of factors in emotional expression, the study further analyzed the effects of environmental stimuli, movement direction, and repetition on emotional expression. Analysis results revealed significant differences in valence recognition due to positive and negative stimuli, vertical repetition, and the interaction between vertical repetition and movement direction. This is because: 1) The vertical and horizontal movement directions have similar effects in expressing valence. Forward and jumping express positive valence, while backward and downward movement express negative valence. 2) When the two directions are mutually exclusive in valence expression, the environmental stimuli has a guiding effect on the recognition results. Under the prompts and guidance of different types of stimuli, participants tend to focus on different types of actions and choose different valence results. Therefore, under positive stimuli, the results tend to be biased towards positive valence. Under negative stimuli, the results tend to be biased towards negative valence.

Additionally, the interaction effect between vertical repetition and motion direction groups significantly influenced recognition results. However, no significant recognition was observed in the downward movement and approach groups ( $D_M \oplus A_p$ and $D_M \otimes A_p$ ). Upon analysis, firstly, we found that under positive stimuli, the approach motion was associated with positive emotions and was more likely to capture attention. Conversely, downward movements, including repeated downward motions, were likely to be ignored as conveying negative expressions. Thus, influenced by positive stimulus scenarios, observers' subconscious neglect for vertical repeating motion caused no significant difference in recognition results. Secondly, under negative stimuli, the emotion type with significant recognition results in $N:D_M \oplus A_p$ was B (+P-A-D): docile; the types of emotions with significant recognition results in group $N:D_M \otimes A_p$ was F (-P-A+D): disdainful/reproach. As the participants were only presented with options of emotional types without specific valence dimensions during the decision-making process, they tended to observe all options and select a consistent emotional attitude. Therefore, by comparing the results of B (+P-A-D): docile and F (-P-A+D): disdainful/reproach, it can be found that repetitive downward movement resulted in a certain degree of decrease in pleasure, but the difference had not yet reached a significant level. Moreover, repetition downward changed the robot from displaying obedience to expressing blame, and its dominance (D) increased from weak to strong. Previous studies have found that the straightness of the spine is seen as a display of prestige [37]. Vertical movement, to some extent, signifies the straightness of the spine. A single downward movement implies spinal curvature, reminiscent of obedience and displaying a submissive stance. Conversely, multiple downward movements necessitate a continuous cycle of returning upward to the initial position and repeating the downward movement. This leads to a visual repetition of upward and downward movement in the vertical direction, emphasizing the vertical trajectory to some extent. This emphasis on the path suggests the straightness of the spine and enhances the expression of emotional dominance.

Secondly, the research results on emotional content selection verified $H_3$: by utilizing holistic factors, non-humanoid robots can effectively convey comprehensible emotional categories to humans. Only $N:U_M \otimes A_v$ cannot be recognized as a relatively unified emotional type. From the previous research results, it can be seen that both stimulation and vertical repetition have a significant impact on the recognition of pleasure. Therefore, under negative stimuli, the positive valence conveyed by vertical repetition conflicts with the negative valence emphasized by negative stimuli, making it challenging for participants to determine a unified and precise emotional result. This happens because our study was designed to discuss the effects of different factors. Emotional expression in real emotional expression environments should be more unified on the expression, and future robot emotional expression should minimize conflicting expressions to enhance user recognition.

Thirdly, detailed research outcomes are outlined in Table VII. After conducting a comprehensive analysis, the following conclusions can be drawn:

TABLE VII. RECOGNITION RESULTS

| stimuli | Vertical movement | Horizontal movement | Valence | Effect of vertical repeated on valence | Emotional results | *Examples* |
|---|---|---|---|---|---|---|
| +[a] | + | − | + | significant increase | +P-A+D | $P:U_M \oplus A_v$ |
| + | Multiple + | − | + | | +P+A+D | $P:U_M \otimes A_v$ |
| + | − | + | + | no significant change | +P-A+D | $P:D_M \oplus A_p$ |
| + | Multiple − | + | + | | +P-A+D | $P:D_M \otimes A_p$ |
| − | + | − | − | significant increase | -P+A-D | $N:U_M \oplus A_v$ |
| − | Multiple + | − | − | | / | $N:U_M \otimes A_v$ |
| − | − | + | − | no significant change | +P-A-D | $N:D_M \oplus A_p$ |
| − | Multiple − | + | − | | -P-A+D | $N:D_M \otimes A_p$ |

- The recognition of robot valence and emotions by users is collectively influenced by various factors, including stimuli, vertical repetition, and motion direction.
- Confusion and identification difficulties can arise when expressive factors contradict the intended emotional expression. Future design should aim to avoid situations where factors are mutually exclusive.
- Under the influence of stimuli, the combination of vertical and horizontal movements helps to intuitively display complex emotions by demonstrating valence and a certain degree of emotional dominance. However, further research is needed on the factors that affect the expression of dominance.

## V. GENERAL DISCUSSION

Through subjective measurement methods, we found the potential rules that worked in humans recognizing emotional actions. Through experiments, we confirmed the role of holistic factors in the hypothesis. Meanwhile, this study validated the effect of Gestalt theory on emotion recognition.

The value of holistic expression factors lies in finding decisive expression elements from complex factors, which to some extent reduces the difficulty of expression for non-humanoid robots and achieves simplification of complex problems. For robots with a high degree of anthropomorphism, there are many factors that can help express emotions, but the effects of these factors are random. When faced with different types of robot expression tasks, designers often find it difficult to judge and select truly valuable expression elements, and the quality of expression results is also difficult to predict. Therefore, the analysis of universal holistic factors helps us understand the key elements that truly play a role in expression, allowing designers to more efficiently select and apply relevant expression materials for design. At the same time, these holistic elements are also important feature vectors for various types of robots in the future to achieve autonomous expression using algorithms. As a result, this study provides new ideas for the future development of emotion expression in non-humanoid robots.

## VI. CONCLUSION

The growing integration of non-humanoid robots into social life has elevated the significance of researching their capabilities of emotional expression. In this study, we introduced holistic expression factors and theoretical frameworks based on Gestalt theory and SOR theory. Subsequently, we performed experiments to validate the function of these holistic expression factors and their influence on emotional expression.

In the first experiment, we confirmed the significance of horizontal and vertical movement directions in conveying emotional valence. In the second experiment, we examined the influence of stimuli, movement direction, repetition, and their collective impact on emotional expression as holistic expression factors. When a non-humanoid robot can move both vertically and horizontally, there is no dominant direction influencing valence recognition results. Recognition results are affected by environmental stimuli, vertical repetitive movements, and the interplay of factors. The results indicate that horizontal and vertical motion expressions can influence the manifestation of valence and emotional dominance based on environmental stimuli, thereby helping non-humanoid robots to present emotionally recognizable content to humans.

These results validate that human perception of emotional expression actions is also influenced by a holistic perspective, introducing a novel viewpoint into the process of designing emotional motions for robots. In consequence, it shifts the conventional understanding among robot emotion expression designers, emphasizing that robots can convey emotions beyond mere imitation of specific human or organismic characteristics. Adopting a holistic approach enables various types of non-humanoid robots to improve the efficiency of their expression.

Future research can investigate the holistic expression factors related to emotional arousal and emotional dominance, enabling non-humanoid robots to convey a wider range of emotions.

## REFERENCES


[1] Bonarini, "Can my robotic home cleaner be happy? Issues about emotional expression in non-bio-inspired robots," Adaptive Behavior, 2016.

[2] T. Law, J. D. Leeuw, J. H. Long, "How Movements of a Non-Humanoid Robot Affect Emotional Perceptions and Trust," International Journal of Social Robotics, 2020.

[3] M. Faria, A. Costigliola, P. Alves-oliveira, et al., "Follow me: Communicating intentions with a spherical robot," proceedings of the 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), 2016 .

[4] G. I. Na, L. PA, P. Keepon, et al., "A playful robot for research, therapy, and entertainment," Intjsocial Robotics, 2009.

[5] M. Szabóová, M. Sarnovsky, V. Kreňáková, et al., "Emotion Analysis in Human–Robot Interaction," Electronics, 9(11): 1761, 2020.

[6] J. R. Cauchard, K. Y. Zhai, M. Spadafora, et al., "Emotion Encoding in Human-Drone Interaction," The Eleventh ACM/IEEE International Conference on Human Robot Interaction. Christchurch, New Zealand; IEEE Press, pp. 263–270, 2016.

[7] G.Venture, D. Kuli, "Robot Expressive Motions: A Survey of Generation and Evaluation Methods," proceedings of the ACMPUB27New York, NY, USA, 2019.

[8] T. R. Groechel, Z. Shi, R. Pakkar, et al., "Using Socially Expressive Mixed Reality Arms for Enhancing Low-Expressivity Robots ," IEEE, 2019.

[9] A. Valenti, A. Block, M. Chita-tegmark, et al., "Emotion Expression in a Socially Assistive Robot for Persons with Parkinson's Disease," 2020.

[10] D. O. Johnson, R. H. Cuijpers, "Investigating the Effect of a Humanoid Robot's Head Position on Imitating Human Emotions," International Journal of Social Robotics, 2017.

[11] R. Savery, R. Rose, G. Weinberg, "Establishing Human-Robot Trust through Music-Driven Robotic Emotion Prosody and Gesture," IEEE, 2020.

[12] T. Ribeiro, A. Paiva. "Nutty-based Robot Animation -- Principles and Practices," 2019

[13] D. Szafir, B. Mutlu, T. Fong, "Communication of Intent in Assistive Free Flyers," proceedings of the Acm/ieee International Conference on Human-robot Interaction, 2014.

[14] J. Forlizzi, C. Bartneck, J. Yamato, et al., Bringing Robots To Life: Applying Principles Of Animation To Robots. 2004.

[15] R, Laban, L. Ullmann, "The Mastery of Movement," Creative Activities, pp: 200, 1971.

[16] S. J. Burton, A. A. Samadani, R. Gorbet, et al., "Laban Movement Analysis and Affective Movement Generation for Robots and Other Near-Living Creatures," Springer International Publishing, 2016.

[17] R. Urtasun, P. Glardon, R. Boulic, et al., "Style-ased Motion Synthesis," Computer Graphics Forum, 23(4): 799-812, 2004.

[18] O. Alemi, W. Li, P. Pasquier, "Affect-expressive movement generation with factored conditional Restricted Boltzmann Machines," proceedings of the International Conference on Affective Computing & Intelligent Interaction, 2015 .

[19] Z. Wang, K. Muelling, M. P. Deisenroth, et al., "Probabilistic movement modeling for intention inference in human–robot interaction," International Journal of Robotics Research, 2013.

[20] H. J. Smith, C. Cao, M. Neff, et al., "Efficient Neural Networks for Real-time Motion Style Transfer," Proc ACM Comput Graph Interact Tech, 2(2): Article 13, 2019.

[21] S. Hou, W. Xu, J. Chai, et al., "A Causal Convolutional Neural Network for Motion Modeling and Synthesis," 2021.

[22] G. Nie, Y. Zhan, "A Review of Affective Generation Models," arXiv: 2202.10763, 2022.

[23] H. Du, E. Herrmann, J. Sprenger, et al., "Stylistic Locomotion Modeling and Synthesis using Variational Generative Models," Motion, Interaction and Games, 2019.

[24] W. Kohler, "Gestalt Psychology," The New American Library, 1947.

[25] A. Mehrabian, J. A. Russell, An approach to environmental psychology. Cambridge: MIT Press, 1974.

[26] A. J. Elliot, A. B. Eder, Harmon-Jones E. "Approach-Avoidance Motivation and Emotion: Convergence and Divergence," Emotion Review, 5(3): 308-11, 2013.

[27] J. D. Laird, K. Lacasse, "Bodily Influences on Emotional Feelings: Accumulating Evidence and Extensions of William James's Theory of Emotion," Emotion Review, 6(1): 27-34, 2014.

[28] J. A. Russell, "A Circumplex Model of Affect," Journal of Personality and Social Psychology, 39(6): 1161-78, 1980.

[29] S. H. Paplu, C. Mishra, K. Berns, "Real-time Emotion Appraisal with Circumplex Model for Human-Robot Interaction," arXiv, 2022.

[30] N. Spatola, O. A. Wudarczyk, "Implicit Attitudes Towards Robots Predict Explicit Attitudes, Semantic Distance Between Robots and Humans, Anthropomorphism, and Prosocial Behavior: From Attitudes to Human–Robot Interaction," International Journal of Social Robotics, pp: 1-11, 2020.

[31] I. Leite, C. Martinho, A. Pereira, et al., "iCat: an Affective Game Buddy Based on Anticipatory Mechanisms," proceedings of the 7th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS 2008), Estoril, Portugal, May 12-16, Volume 3, 2008 .

[32] M. M. Bradley, P. J. Lang, "Measuring emotion: the Self-Assessment Manikin and the Semantic Differential," Journal of behavior therapy and experimental psychiatry, 25(1): 49-59, 1994.

[33] J. Xu, J. Broekens, K. Hindriks, et al., "Mood contagion of robot body language in human robot interaction," Autonomous Agents & Multi-Agent Systems, 2015.

[34] J. Elster, "Strong feelings: emotion, addiction, and human behavior," Jean Nicod Lectures, 2000.

[35] R. Plutchik, The nature of emotions: Human emotions have deep evolutionary roots, 2001.

[36] P. Gebhard, "ALMA: a layered model of affect," proceedings of the 4th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS 2005), July 25-29, Utrecht, The Netherlands, 2005.

[37] R. Krner, A. Schütz, "Dominance or prestige: A review of the effects of power poses and other body postures," Social and Personality Psychology Compass, 2020.