

Received 24 December 2024, accepted 20 January 2025, date of publication 27 January 2025, date of current version 3 February 2025.

Digital Object Identifier 10.1109/ACCESS.2025.3534845



RESEARCH ARTICLE

The Development of Human–Robot Interaction Design for Optimal Emotional Expression in Social Robots Used by Older People: Design of Robot Facial Expressions and Gestures

SUJIN JO^{ID} AND SEONGSOO HONG^{ID}

Department of Design Engineering, Tech University of Korea, Siheung 15073, South Korea

Corresponding author: SeongSoo Hong (hss@tukorea.ac.kr)

This work was supported by the Ministry of Trade, Industry and Energy, South Korea, under the project “Development of Design Industry Technology: Building Capacity for Design Innovation” (Project No. 20009760: Development of human friendly multipurpose service robot and new market creation by applying human robot interaction design).

This work involved human subjects or animals in its research. The authors confirm that all human/animal subject research procedures and protocols are exempt from review board approval.

ABSTRACT Showing facial expressions and using emotion-appropriate gestures are essential for social robots. As a robot's behavior becomes more anthropomorphic, the intimacy and naturalness of human-robot interactions improve. This study aims to derive optimized facial expression and gesture designs for social robots interacting with elderly individuals, thereby enhancing emotional interactions. First, we utilized user-robot integrated scenarios to identify the emotional states required for robot interactions. Subsequently, we conducted surveys and user preference evaluations on commercially available robot faces. The results indicated that suitable components for robot faces include the eyes, eyebrows, mouth, and cheeks; geometric shapes were deemed the most appropriate. Accordingly, we collected and analyzed human facial expression images using the Facial Action Coding System to identify action unit combinations and facial landmarks. This analysis informed the design of robot faces capable of expressing humanlike emotions. Furthermore, we collected and evaluated human gesture videos representing various emotions to select the most suitable gestures, which were analyzed using motion capture technology. We utilized these data to design robot gestures. The designed robot facial expressions and gestures were validated and refined through emotion-based user preference evaluations. As a result of the study, we developed facial expression and gesture designs for six emotions (Loving, Joyful, Upbeat, Hopeful, Concerned, Grateful) in social robots interacting with elderly individuals. The results provide guidelines for designing human-friendly robot facial expressions and gestures, thus enabling social robots to form deep emotional bonds with users. By analyzing human facial expressions and gestures in relation to emotions and applying these findings to robots, we successfully developed natural and emotionally expressive robot behaviors. These findings contribute to the advancement of robots as reliable and comforting companions for humans.

INDEX TERMS Human–robot interaction (HRI) design, robot facial expression design, robot gesture design, social robot design.

I. INTRODUCTION

A. STUDY BACKGROUND AND PURPOSE

A social robot is defined as a robot that is capable of engaging in emotional and social interactions with humans through

The associate editor coordinating the review of this manuscript and approving it for publication was Byoung Wook Choi^{ID}.

the utilization of social behaviors, including conversations, facial expressions, and gestures [1]. For a social robot to interact effectively with humans, it must be able to comprehend human emotions and intentions and respond appropriately. Facial expressions and gestures play a significant role in this process. As posited by Li et al. [2], humans tend to exhibit a heightened sense of proximity and affinity toward robots

that exhibit behaviors analogous to those observed in humans, including gestures, facial expressions, and emotional states. In other words, if a robot exhibits behavior that is perceived as similar to human behavior, users will experience greater comfort when interacting with it [2].

However, the facial expressions and gestures of modern social robots are often simplistic and mechanical in nature, hindering their ability to convey emotions accurately. Moreover, as robot facial expressions and gestures are expressed differently from those of humans, users may find it difficult to understand these emotional expressions and interact effectively with such robots.

The objective of this study is to develop an optimized set of robot facial expressions and gestures that can enable robots to express emotions in a manner similar to humans. Through this, we aim to facilitate robot interactions that foster natural emotional connections with users.

The findings of this study provide guidelines for effectively applying facial expressions and gestures based on the emotions that social robots should convey during the commercialization phase. These contributions aim to advance user-centered Human-Robot Interaction (HRI) design.

B. STUDY SCOPE AND METHODS

This study focuses on social robots that interact with elderly individuals and provide them with emotional care through affective support. For social robots to effectively interact with users, the implementation of emotional expression and human-robot interaction (HRI) is essential. Therefore, we aim to develop interaction designs based on robot emotions.

This research involved the following steps: Initially, we conducted a theoretical review of the definition and principal elements of HRI design. Subsequently, we devised a series of integration scenarios in which elderly users interact with a social robot. Based on these scenarios, we derived the emotions that the robot should express.

In particular, we analyzed existing research on the relationship between emotions and facial expressions. Prior research has demonstrated that a one-to-one correspondence between emotions and facial expressions is an inaccurate assumption, as facial expressions can vary significantly between individuals and as the same emotion can be expressed in various ways. Moreover, it has been argued that facial expressions are not an exact representation of emotions [3], [4]. However, in the context of social robots, it is crucial that facial expressions provide contextually consistent nonverbal cues to facilitate users' comprehension of a robot's state and intentions. This represents a distinct approach to the interpretation of emotions conveyed through facial expressions in human interactions.

Accordingly, in this study, we propose a design that streamlines and standardizes the facial expressions of social robots, thereby enabling the clear and intuitive conveyance of emotions. This facilitates the formation of an emotional bond

between a robot and its user and thus fosters more friendly and natural interactions.

The following section details the utilization of Ekman's Facial Action Coding System (FACS) to analyze human facial expressions and subsequently design robot facial expressions based on these observations. To address the limitations of Ekman's theory, we examined Barrett's theory in parallel and conducted additional research on emotional gestures. The robot in this study was designed to express emotions using both facial expressions and gestures, and we confirmed the efficacy of this approach through a user verification evaluation.

Specifically, we collected images of human facial expressions representing different emotions. Based on these images, we created average facial expression images for each emotion. We then analyzed the facial expressions in the average images using the FACS and facial expression landmarks. The robot's facial expressions were designed based on the features of the analyzed facial expressions.

In the gesture design process, we collected videos of human emotional gestures and conducted preference evaluations to select the optimal gesture videos that best represented emotions. The selected videos were motion captured to establish key poses and derive kinematic elements, which were used to design the robot's gestures. During this process, we optimized the robot's gesture data using three methods: joint data matching, kinematic factor analysis, and additional compensation.

Subsequently, we conducted a preference evaluation to ascertain the suitability of the designed robot's facial expressions and gestures. Accordingly, we finalized the design such that these facial expressions and gestures would facilitate an emotional connection and enable the clear conveyance of emotions in interactions with users.

In this study, we propose an interaction design for a social robot that can support emotional care in interactions with elderly users through HRI design elements. The design is centered on integrated facial expressions and gestures, which are designed according to emotional expressions. Furthermore, by integrating Ekman's and Barrett's theories, we develop a novel approach to robots' emotional expression, supported by scientific evidence derived from an integrated design of facial expressions and gestures.

II. UNDERSTANDING HRI DESIGN

Hong [5] defined the process of HRI design as the method by which humans provide commands or receive information from robots. He further elucidated that HRI design is the concept of facilitating interactions with robots so that users can perform desired functions and posited that HRI design is an instrumental factor in shaping the relationship between a robot and its user. He also argued that the key to HRI design is enabling interactions that integrally coordinate a user's needs [5].

In a recent study, Lee and Kim [6] proposed a classification of the elements of nonverbal HRI. This classification is based

on four main categories: visual elements (text, shape, and color), auditory elements (volume, timbre, and sound speed), movement elements (motion pattern and speed), and touch elements (temperature, material, and vibration) [6]. Soo et al. additionally posited that gestures and facial expressions are crucial elements in the hardware aspect of robots, whereas sounds are essential in the software aspect to indicate a robot's state [7].

Considering the aforementioned details, we have organized the principal elements of HRI design into the following categories, as shown in Table 1: the gestures, facial expressions, appearance, and sound of a robot. The gesture element pertains to the emotional behavior of a robot, whereas the facial expression element concerns how a robot expresses emotions using its face. A robot's appearance is constituted by several characteristics, including color, shape, material, and temperature. The sound element encompasses dialogue, sounds, tones, and effects.

TABLE 1. HRI design elements.

HRI Design Elements	Details
Gestures	Behavior patterns, speed of motion, etc.
Facial expressions	Graphics, such as symbolized emoticons, colors, characters, etc., on the display
Appearance	Color, shape, material, temperature, proportion, size, weight, etc.
Sound	Voice, text-to-speech (TTS), acoustics, pitch, etc.

III. ELICITING EMOTION INFORMATION FROM SOCIAL ROBOTS FOR THE ELDERLY

A. USER-ROBOT INTEGRATION SCENARIOS

The scenario-based design process is a methodology that is concerned with discovering user needs through the presentation of specific usage situations [8]. The user-robot integration scenario presented in this study is an integrated scenario based on the aforementioned scenario. Our integrated scenario serves to elucidate the contextual interactions from the perspectives of both the user and the robot. It may be employed to gain insights into the interactions between the three elements involved: the robot, the user, and the environment.

In this study, we constructed an integrated scenario of the interactions between social robots and humans featuring the elderly, from the time they wake up in the morning to the time they go to bed, as shown in Table 2.

B. ELICITING EMOTION INFORMATION FROM SOCIAL ROBOTS

It is essential that a robot is able to express appropriate emotions to meet its users' needs and respond to requests. To achieve this, we conducted group interviews using a mood meter, with the aim of eliciting appropriate robot emotions for

TABLE 2. User-robot interaction scenarios (*U-users, *R-robot).

Situation	Interaction Images	User-Robot Interaction	
Greetings		*U	The elderly man says good morning to the robot. He talks about his mood and physical condition.
		*R	The robot greets the elderly as they wake up time. The robot talks to the elderly to check on their conditions.
Today's news		U	The elderly learn about things such as the weather from the robot.
		R	The robot informs the users about the day, such as the weather and top news.
Breakfast		U	The elderly man enjoys talking to the robot while eating.
		R	The robot accompanies the elderly during mealtime. It talks to them about the flavor of the food, their interests, and more.
Medication reminders		U	The elderly man forgets to take his medication but corrects this when the robot reminds him to take it.
		R	The robot reminds the elderly at the right time to take their medication. The robot asks them if they have taken their medication and then offers it to them.
Take your medication		U	The elderly person takes their medication.
		R	The robot praises the elderly person for taking their medication.
Game suggestions		U	The elderly play dementia-prevention games with the robot's suggestions.
		R	The robot suggests dementia-prevention games to bored users.
Playing games		U	One elderly man finds the game difficult but quickly gets into it and interacts with the robot.
		R	The robot praises the elderly's correct answers, encourages them, and cheers them on.
Encouraging exercise		U	The elderly man is sitting down, and he starts exercising when the robot suggests it.
		R	The robot reaches out to sedentary users and encourages them to exercise and stretch.
Workout progression		U	The elderly feel safe and supported by exercising with the robot.
		R	The robot guides the elderly through the movements several times, offering encouragement and praise.
Emergencies		U	The elderly man falls while exercising. He is surprised but reassured by the robot's checks and care.
		R	The robot detects that an emergency has occurred. It checks the person's status and waits for an emergency call.
Bedtime greetings		U	The elderly man says good night to the robot at the end of the day and goes to sleep.
		R	The robot checks in on the elderly person's day and says good evening.

the scenario. The mood meter is a tool that was introduced in Brackett's book, *Permission to Feel* [9].

Ten elderly individuals (four men and six women) aged 65 years or older—which is the age range of the intended users of our robot—participated in the group interviews. The participants were divided into two groups, designated as Group A and Group B, with five participants each. The objective was to identify user-robot integration scenarios and derive robot emotions for each interaction. Based on the mood meter and via group discussions, the most appropriate robot emotions for each robot interaction were selected using a step-by-step process, as shown in Table 3.

TABLE 3. Emotions derived from group interviews.

Situation	Robot Interaction	Emotions Elicited by Groups	
Greetings	The robot greets the elderly as they wake up time. The robot talks to the elderly to check on their conditions.	Group A	Peaceful
		Group B	Loving
Today's news	The robot informs the users about the day, such as the weather and top news.	Group A	Loving
		Group B	Balanced
Breakfast	The robot accompanies the elderly during mealtime. It talks to them about the flavor of the food, their interests, and more.	Group A	Relaxed
		Group B	Loving
Medication reminders	The robot reminds the elderly at the right time to take their medication. The robot asks them if they have taken their medication and then offers it to them.	Group A	Loving
		Group B	Chill
Take your medication	The robot praises the elderly for taking their medication.	Group A	Upbeat
		Group B	Joyful
Game suggestions	The robot suggests dementia-prevention games to bored users.	Group A	Upbeat
		Group B	Playful
Playing games	The robot praises the elderly's correct answers, encourages them, and cheers them on.	Group A	Hopeful
		Group B	Joyful
Encouraging exercise	The robot reaches out to sedentary users and encourages them to exercise and stretch.	Group A	Loving
		Group B	Upbeat
Workout progression	The robot guides users through the movements several times, offering encouragement and praise.	Group A	Hopeful
		Group B	Loving
Emergencies	The robot detects that an emergency has occurred. The robot checks the person's status and waits for an emergency call.	Group A	Concerned
		Group B	Restless
Bedtime greetings	The robot checks in on the elderly person's day and says good evening.	Group A	Grateful
		Group B	Relaxed

To ascertain the suitability of the emotional data obtained from the group interviews, a seven-point questionnaire was administered to ten experts (five men and five women) in the field of robotics. The experts awarded higher scores to the emotions selected by the two groups that were deemed more

appropriate for the situation. The higher-scoring emotions were subsequently identified as the most appropriate for the robot to express in the given context, as shown in Table 4. The selected emotions were “loving,” “joyful,” “upbeat,” “hopeful,” “concerned,” and “grateful.”

TABLE 4. Evaluation results by robotics experts.

Situation	Derived Robot Emotions	Total Expert Review Score	Average
Greetings	Peaceful	43	4.3
	Loving	52	5.2
Today's news	Loving	54	5.4
	Balanced	39	3.9
Breakfast	Relaxed	35	3.5
	Loving	52	5.2
Medication reminders	Loving	51	5.1
	Chill	42	4.2
Take your medication	Upbeat	42	4.2
	Joyful	54	5.4
Game suggestions	Upbeat	52	5.2
	Playful	36	3.6
Playing games	Hopeful	56	5.6
	Joyful	43	4.3
Encouraging exercise	Loving	46	4.6
	Upbeat	51	5.1
Workout progression	Hopeful	55	5.5
	Loving	41	4.1
Emergencies	Concerned	52	5.2
	Restless	39	3.9
Bedtime greetings	Grateful	52	5.2
	Relaxed	45	4.5

IV. DESIGNING FACIAL EXPRESSIONS FOR SOCIAL ROBOTS FOR THE ELDERLY

A. ROBOT FACE DESIGN

For robots to communicate in a natural manner with users, implementing contextually relevant facial expressions is crucial. At present, the majority of commercially available robots express facial expressions through the use of either two-dimensional (2D) displays or three-dimensional (3D) actuators, with a full-face 2D display being the most common.

In this study, we examined over 50 commercially available robot face designs with the objective of designing a social robot that employs these 2D facial expressions. Based on this analysis, we categorized the samples into four types according to facial components, as shown in Table 5.

We conducted a face-type preference evaluation using a seven-point scale with 20 prospective users of the social robot (people aged 65+) based on the established sample. The Type B face design, characterized by geometric eyes and other components, achieved the highest mean score (5.55), as shown in Table 6. The results of the analysis of variance demonstrated that the P-value was less than 0.01 and thus statistically significant.

To select additional components for Type B, we analyzed the facial components of over 50 robots. Our findings

TABLE 5. Robot face types based on facial components.

Type	Robot Face Design Sample Image		Components
Type A			Geometric eyes
Type B			Geometric eyes and other components
Type C			Detailed eyes (sclerae, pupils, light reflection, etc.)
Type D			Detailed eyes and other components

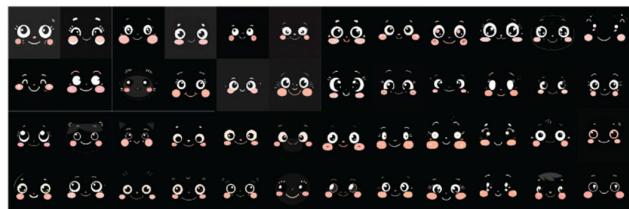
TABLE 6. User preference evaluation results regarding the type of robot faces.

Types	Type A	Type B	Type C	Type D
	Geometric eyes	Geometric eyes and other components	Detailed eyes	Detailed eyes and other components
Scores	4.3	5.55	3.95	3.75

revealed the eyes (50 robots), mouth (32 robots), eyebrows (22 robots), cheeks (15 robots), nose (six robots), hair (three robots), and hands (one robot) to be the most frequently utilized facial components. The eyes are the most important element of emotional expression, as they convey a wide range of emotions. Similarly, the mouth, eyebrows, and cheeks are instrumental in conveying a range of emotions, including joy, anger, and shyness, respectively.

Accordingly, our robot's face was designed with the eyes, mouth, eyebrows, and cheeks in mind, as these components were identified in the findings of over 30% of the robots studied.

To create the face design, we employed MidJourney, an artificial intelligence (AI) tool designed for image generation. This allowed us to gather a multitude of design references, as shown in Fig. 1.

**FIGURE 1.** Collecting face design references through MidJourney. The MidJourney service is designed to facilitate the generation of images characterized by a creative and distinctive style, making it an ideal choice for those seeking artistic inspiration or abstract images [10].

We created six distinct robot face designs, each informed by the design references collated by MidJourney, as shown in

Table 7. Subsequently, we conducted a preference evaluation to select the optimal robot face design based on the aforementioned criteria.

This preference evaluation was conducted using a seven-point scale with 30 prospective users of social robots, comprising individuals aged 65 and above, and the design of Type 3 was identified as the most favorable, with an average score of 4.7, as shown in Table 7. The analysis of variance yielded a P-value that was less than 0.01, which was statistically significant. Consequently, Type 3 was selected as the face design for the social robot in this study.

TABLE 7. Evaluation results of preferences for six types of face designs.

Type	Type 1	Type 2	Type 3
Scores	3.8	3.3	4.7
Type	Type 4	Type 5	Type 6
Scores	4.2	2.6	2.9

B. DERIVATION OF HUMAN FACIAL EXPRESSION IMAGES

To design the facial expressions of our social robot, we employed DALL-E 3, a text-based image generation model developed by OpenAI. We used it to collect images of facial expressions based on human emotions. DALL-E 3 is well suited to the generation of realistic and detailed expressions and is particularly adept at producing images that are realistic in appearance [11]. Multiple images were generated by entering a prompt comprising the word combination “emotion adjective + facial expression + human.”

The images were collated and processed using the Combine Multiple Faces application, which yielded an average facial expression image. In this context, the term “average facial expression image” refers to a composite representation of the typical facial characteristics observed in elderly individuals. Such an image is created by superimposing images based on the facial components (eyes, nose, mouth, eyebrows, face shape, etc.) present in each elderly facial expression image.

For instance, the average facial expression image for the images gathered with the word combination “joyful + expression + human” as a prompt was a smile with the eyes curved into half-moon shapes and the mouth wide open, as shown in Fig. 2. Similarly, average facial expression images were created for the remaining five robot emotions.

C. ROBOT FACIAL EXPRESSION DESIGN

In this study, we aimed to derive action unit (AU) combinations for average facial expression images and to utilize facial



FIGURE 2. “Joyful” elderly facial expressions and their average image.

expression landmarks to design robot facial expressions with precision.

The FACS is an anatomical system that analyzes the movements of the face in response to the display of emotions. The FACS represents facial muscle movements in terms of AUs, with each AU representing a specific region and the strength of the facial muscles involved. The current FACS comprises 44 Action Units (AUs) designed to analyze a wide range of facial expressions. Among them, 30 focus on primary facial muscle movements, while the remaining units account for auxiliary movements, such as head and eye position changes. The analysis of emotional facial expressions is facilitated by the combination of several AUs, which enable a detailed examination of the facial muscles involved in such expressions [12].

Subsequently, we derived facial landmarks using the Face Landmarking feature provided by Megvii’s Face++ platform. In this process, we analyzed facial landmarks using the 68 facial landmark model developed by Davis E. King, the creator of the Dlib library. These 68 landmarks are distributed as follows: jawline (0-16), eyebrows (17-26), nose (27-35), eyes (36-47), and mouth (48-67).

In this study, points for the eyes (36-47), eyebrows (17-26), and the inside of the mouth (60-67) were selected to correspond with the robot’s facial components. In particular, the landmark points located within the oral cavity were deemed most suitable for incorporation into robot faces. The 30 facial landmark points used in this study are shown in Fig. 3. As the cheeks may impede the accurate recognition of emotions, we have designated certain cheek landmarks for use exclusively when conveying intense emotional states. For instance, when the eye landmark points become narrower and AU 6 is strongly expressed, resulting in the elevation of the cheeks, the cheeks are complementarily expressed to indicate the presence of strong emotions.

We designed the facial expressions of our social robot based on AU combinations and derived facial landmarks. First, we extracted AU combinations from average facial expression images, after which we extracted facial landmarks for each facial component and generated characteristic lines. Subsequently, the AU combinations, facial landmarks, and characteristic lines were matched to the robot’s face to finalize the design of the robot’s facial expressions. Fig. 4

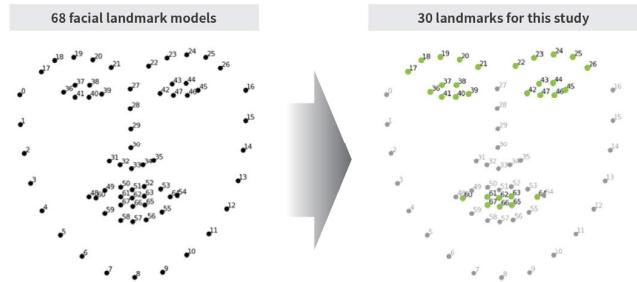


FIGURE 3. Selection of facial landmarks to be applied in the study.

illustrates the process used to derive the robot’s facial expressions based on this methodology.

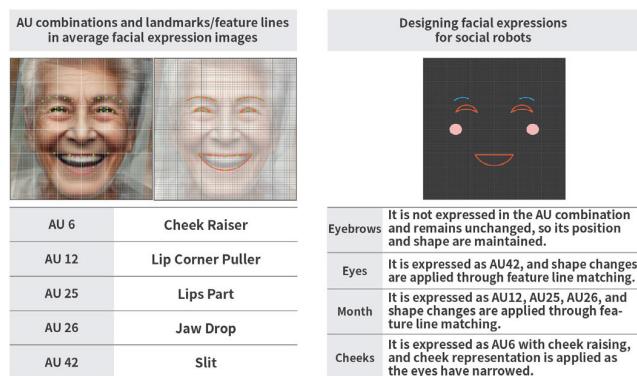


FIGURE 4. Analysis and matching of the “joyful” human expression. The average facial expression image associated with the emotion “joyful” was analyzed and matched to the robot’s facial expression.

The same methodology was employed for the analysis and matching of facial expressions associated with the remaining five emotions, as shown in Table 8.

V. APPEARANCE DESIGN OF SOCIAL ROBOTS FOR THE ELDERLY

A. DESIGNING THE ROBOT’S APPEARANCE

The exterior design, which directly influences the gesture design, was prioritized. This process entailed an investigation of the extant state of robot design, the formulation of a design concept, the delineation of ideas and details, and the ultimate realization of the design.

The initial stage involved analyzing the images of commercially available social robots. This analysis demonstrated that social robots are typically distinguished by their diminutive stature and amiable demeanor, featuring a gentle, curved design that ensures physical safety and fosters a sense of familiarity with users. Furthermore, these robots typically interact with users via simple gestures, including arm, head, and torso movements, with the objective of strengthening emotional bonds. Based on the aforementioned analyses, we derived and sketched a design concept for our social robot, featuring soft curves and fluid lines, with components for the face, torso, and arms. The final design features a wide, rounded face shape, which is conducive to emotional

TABLE 8. Designing robot facial expressions for six emotions in our social robot.

Emotions	Loving	Joyful	Upbeat
Human facial expression analysis			
Derivation of robot facial expressions			
Emotions	Hopeful	Concerned	Grateful
Human facial expression analysis			
Derivation of robot facial expressions			

expression and communication. In addition, the design is intentionally minimalistic, allowing the robot to blend harmoniously into a variety of environments.

B. SETTING THE ROBOT'S RANGE OF MOTION

To ensure the social robot's functionality and safety, the range of motion for each component was carefully calibrated to prevent interference between parts.

The neck is endowed with a single degree of freedom along the x-axis, enabling it to descend by up to 10° in the pitch direction; this allows the head to nod. The arms are endowed with two degrees of freedom along the x-axis and the y-axis, encompassing 120° of forward movement and 60° of backward movement in the pitch direction, as well as 40° of movement in the roll direction. The torso is capable of a single degree of freedom around the z-axis, allowing for a full 360° rotation in the yaw direction. This range is optimal for the expression of emotions without interference from other bodily systems, as shown in Fig. 5.

VI. DESIGNING GESTURES FOR SOCIAL ROBOTS FOR THE ELDERLY

A. DERIVATION OF OPTIMAL GESTURE VIDEOS FOR EACH EMOTION

To develop socially responsive robot gestures that evoke emotional responses, it is essential to collect and analyze human gesture videos to inform the design of robot gestures. At present, robot gestures are characterized by a lack of expressivity and a mechanical quality; this impedes the establishment of a natural rapport with humans. To address this issue, it has been proposed that robots should be programmed

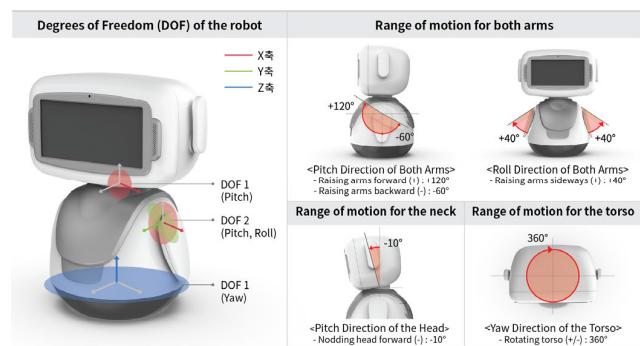


FIGURE 5. Visualization of the operational range of our social robot.

to emulate human behavior, thereby facilitating more natural and convenient interactions.

As posited by Aqdus et al. upper body movements play a greater role than lower body movements in conveying emotions [13]. Accordingly, in this study, we focused on upper body movements. We collected and analyzed videos of human gestures related to emotions, with the objective of reflecting the findings in the design of robot gestures.

We employed a text-based video generation AI, Pictory, to collect gesture videos for each emotion. The term “emotion + person” was entered into Pictory AI to obtain a minimum of five gesture videos for each of the six emotions. For instance, to obtain human gesture videos pertaining to the emotion of “loving,” we entered the text “a loving person” into the search field and collected the related videos.

To derive the optimal gesture video based on the collected human gesture videos, we conducted a preference evaluation using a seven-point scale with 20 prospective users. For instance, Video 3 was rated the highest for the “loving” emotion, with an average score of 6.05. The analysis of variance demonstrated that the P-value was less than 0.01, indicating a statistically significant result.

Similarly, human gesture videos were collected for each of the remaining five emotions, with optimal gesture videos selected through a preference evaluation process, as shown in Table 9.

B. REVIEW OF PRIOR RESEARCH

To inform the design of robot gestures, we investigated previous studies that had applied human gestures to robots. To design gestures for educational robots, Jeong and Hong (2022) codified human upper body movements and established two principles that could be applied to robots. These principles were subsequently used to inform the design of robot gestures [14]. Jung and Hong [15] employed motion capture data to derive motion-based HRI for an unmanned café robot. A scenario was devised based on the interaction between a barista and a user, and the motion capture data were subjected to analysis and subsequently applied to the robot [15]. Song et al. [16] employed motion capture tech-

TABLE 9. Video of optimal gestures for six emotions.

Robot Emotions	Optimal Gesture Video	Evaluation Results	
		Sum	Sig
Loving		6.05	0.00
Joyful		5.85	0.00
Upbeat		6.1	0.00
Hopeful		5.25	0.00
Concerned		5.85	0.00
Grateful		5.95	0.00

nology to assist humanoid robots in accurately replicating human movements. The data were converted to align with the robot's joint structure and range of motion, and optimized movements were implemented [16]. Using a motion capture database, Kim et al. [17] investigated the most effective method for accurately reproducing humanlike upper body motion in a humanoid robot. The data were converted and preprocessed for the robot, enabling the robot to learn and reproduce a range of human upper body movements [17]. Considering these findings, in this study, we aimed to capture and analyze the optimal gesture video for each emotion, convert the data to align with the robot's joint structure and range of motion, and apply these data to the robot.

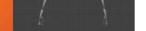
C. MOTION CAPTURE OF OPTIMAL GESTURE VIDEOS

The utilization of AI-based motion capture technology is an optimal methodology for the extraction of motion data from optimal gesture footage. It is challenging to obtain precise data if the most suitable gesture video is rerecorded, as the outcomes may fluctuate depending on the performer. In this study, we employed DeepMotion's AI-based motion capture technology to generate motion data from optimal gesture videos. This technology tracks and analyzes movement in videos to generate 3D skeleton data, offering a cost-effective method for obtaining high-quality motion data without complex equipment.

Nevertheless, it should be noted that AI-based motion capture technology is susceptible to errors in instances of complex or fast motion, as well as in cases where the subject is partially obscured. To enhance this process, we derived a "key pose," which denotes a crucial motion. In a 2010 publication, Yamane asserted that the key frame method represents the most effective approach for extracting pivotal motion points from human motion capture data and applying them to characters [18]. In this study, the concept of a key frame has been defined as a "key pose."

Deriving key poses enabled us to analyze the motion capture data on a key-pose-by-key-pose basis, with the aim of applying the findings to our robot in a manner as close as possible to human gestures. To illustrate this, Table 10 provides a summary of the key poses for the "loving" emotion; the remaining five emotions were also captured and keyed in a similar manner.

TABLE 10. Key pose for the emotion "loving."

Key Pose	Images (Video/Motion Capture)	Key Pose Descriptions	Frame (30 fps)
1	 	Raise your right hand and wave	1-44
2	 	Lower your right hand slightly	45-52
3	 	Raise both hands and wave	53-98
4	 	Lower both hands	99-120

D. DERIVATION OF KINEMATIC ELEMENTS FOR GESTURE DESIGN

To apply motion capture data to social robot gestures, it is essential to derive kinematics. These are physical variables, including position, speed, and acceleration, which can be employed in the analysis of motion data to inform the design of humanlike robot gestures.

Several researchers have proposed that Laban Efforts may prove to be a valuable framework for the design of robot behavior [19]. Laban Efforts constitute a component of the Laban movement analysis (LMA), a system devised by Rudolf von Laban in the 1960s for recording dance choreography. Laban [20] posited that Efforts comprise four factors, namely space, flow, time, and weight, that are employed to effectively express movement goals [20], [21].

In a study conducted by Knight and Simmons [19], robot movement was analyzed using LMA. This revealed a lack of consistency in the Efforts factors across existing research, as well as a notable absence of the flow factor. He employed all four Efforts factors to generate robot motion and demonstrated that reliable motion could be achieved even for robots with low degrees of freedom [19].

In this study, we analyzed motion data based on the kinematic components of the Efforts factors, as established by Knight and Simmons [19], and applied them to social robot gestures. Table 11 delineates the kinematic factors employed in our study.

TABLE 11. Kinematic elements applied in this study.

Efforts Factor	Kinematic Factor	Definition
Space	Direction	Direction in which the robot is moving (x-, y-, and z-axis)
Flow	Range	Displacement (in degrees) by which the robot's position moves
Time	Speed	The speed of the robot while it moves the target travel distance
Weight	Acceleration	The amount of speed change while the robot is moving

E. METHOD FOR APPLYING MOTION CAPTURE DATA

As the social robot in this study has a different number of joints and degrees of freedom than humans, it is not possible to directly apply human motion capture data to its gestures. Therefore, based on the derived key poses, we converted only the joint data necessary to express those poses into robot gestures.

To achieve this, we devised three methods for applying motion data, taking into account the robot's joint structure and range of motion: joint data matching, kinematic factor analysis, and additional compensation.

We employed the open-source 3D graphics tool Blender to analyze motion capture data. In the following sections, we refer to the motion data obtained through Blender as "motion data."

1) HOW TO APPLY JOINT DATA MATCHING

The method for applying joint data matching involves selecting the requisite joints from the motion data and then matching them to a robot's joints.

The initial step is to establish a correlation between the requisite joints within the motion capture data and the corresponding joints on a robot. This process entails the exclusion of superfluous joints because, when matching data with robot joints, it is advisable to select those that yield results that are as close as possible to the motion capture data. In addition, only the requisite joint data are converted into robot gestures. Table 12 illustrates the outcome of the process of matching motion data with our robot's joints.

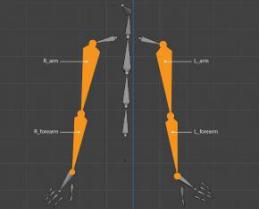
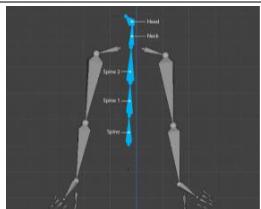
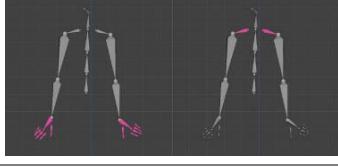
2) HOW TO APPLY KINEMATIC FACTOR ANALYSIS

This section details how to analyze motion capture data and apply the results to a robot.

First, the direction is determined by analyzing the direction of joint motion and assigning it to the axis along which the greatest motion occurs. For instance, if the joint motion in the motion data indicates that the greatest motion occurs in the x-axis direction among the x-, y-, and z-axis, the motion direction of the robot joint is the x-axis. This direction is employed as the reference axis for deriving the acceleration coordinate value graph.

Second, the range in question prioritizes the motion with the largest range of motion of the matched joints. For

TABLE 12. Matching of motion data joints with robot joints.

Joints in Motion Data	Joints in the Robot
	 • Joint A Pitch
Motion data: The motion of the head, neck, spine, Spine 1, and Spine 2 joints in the x-axis direction is matched to the robot's Joint A (pitch).	
	 • Joint B Pitch Joint B Roll
Motion data: The motion of the R_arm, R_forearm, L_arm, and L_forearm of a joint is defined by the robot's Joint B (pitch/roll) along the axis direction.	
	 • Joint C Yaw
Motion data: The motion of the head, neck, spine, Spine 1, and Spine 2 joints in the y- and z-axis directions is matched with the robot's Joint C (yaw).	
Excluded Joints	
	Shoulder joints are excluded as they are not involved in the action, and finger joints are excluded because emotions can be expressed solely through arm movements.

instance, when a robot's right arm gesture is implemented, if R_arm has a rotation range of 80° and R_forearm has a rotation range of 160° , the range of R_forearm—which has the larger rotation range—is applied. Nevertheless, should the rotation exceed the maximum range of the robot, it must be scaled to that range.

Third, the principle of speed factor analysis is applied by calculating the amount of range change divided by the amount of frame change. In this study, the motion of the robot is expressed in terms of the rotation of each component around the robot's joint axis. Accordingly, the speed is calculated as the range change divided by the time change. The time change is calculated as the change in the frame. For instance, in Emotion A, if the range is 0.1 radians in Frame 1 and 3.14 radians in Frame 48 and if the frames per second (FPS) is 30, the calculation is as follows:

$$\text{Range change} : 3.14 - 0.1 = 3.04 \text{ radians}$$

$$\text{Frame change} : \text{Frames/FPS} = (48 - 1)/30 = 47/30$$

Rate : Range change/frame change = approx. 1.94 radians/second

Fourth, the acceleration is reflected by the setting of the motion Bézier curve, which is based on the Euler rotation graph. This graph is a visual representation of the change in the rotation angle over time, which enables a robot to demonstrate acceleration patterns comparable to those observed in the motion data.

Previous research has demonstrated that the acceleration of robot gestures can be derived by plotting the time-varying coordinate values of motion capture data and setting the Bézier curve of the robot motion to have the same shape as the plot [15].

In this study, the objective of utilizing acceleration was not to calculate and apply an exact numerical value of acceleration; rather, we aimed to approximate the pattern of speed changes over time in the motion data to the robot's motion. Although Euler rotation graphs are not acceleration graphs, they provide the most accurate speed change pattern because they show the change in the rotation angle over time. Therefore, we concluded that applying acceleration based on the Euler rotation graph would be the most appropriate method.

For instance, the acceleration of motion data exhibiting substantial movement along the x-axis can be applied to robot gestures. This can be achieved by deriving a pattern of angular change based on the x-axis Euler rotation graph. The pattern derived from this graph should be applied to the acceleration graph of robot gestures, with the objective of achieving an acceleration pattern similar to that observed in the motion data.

3) HOW TO APPLY ADDITIONAL COMPENSATION

It should be noted that, due to the inherent limitations of the motion capture technique, the results of the motion capture process may not fully reflect the behavior expressed in the optimal human gesture footage and key poses. To address this issue, the following process can be undertaken.

It is imperative that all movements both commence and conclude in a standing position, before progressing to the subsequent key pose. In the event that this is not the case, the angles must be adjusted to compensate for this discrepancy, ensuring that the initial and final movements are performed in a standing position. Furthermore, the acceleration graph of the final movement should be modified to exhibit a downward curve, thereby facilitating a natural transition to the subsequent movement.

Furthermore, in instances in which a gesture is included in the key pose but not reflected in the motion capture, the optimal gesture video can be referenced to supplement the representation. For instance, if a hand wave is not captured, the video can be used to compensate for this omission. In addition, when both arms perform the same action, their range, direction, speed, and acceleration are aligned. This process unifies the subtle differences in human motion, enabling a robot to move both arms in a consistent manner.

F. STRUCTURE OF MOTION CAPTURE DATA

In this study, we analyzed motion capture data using Blender. Upon importing the motion capture data as a BVH file, a skeletal structure comprising multiple bones was revealed, as illustrated in Fig. 6. Furthermore, the animation frames (which contain temporal data) and channels (which contain movement data) can be observed. A bone comprises a head that corresponds to the rotation axis and a tail that corresponds to the endpoint. The channel contains the displacement values of the head and the tail. The displacement value is subject to variation as a consequence of changes in the frame [15]. The motion capture data employed in this study were extracted at a rate of 30 FPS.

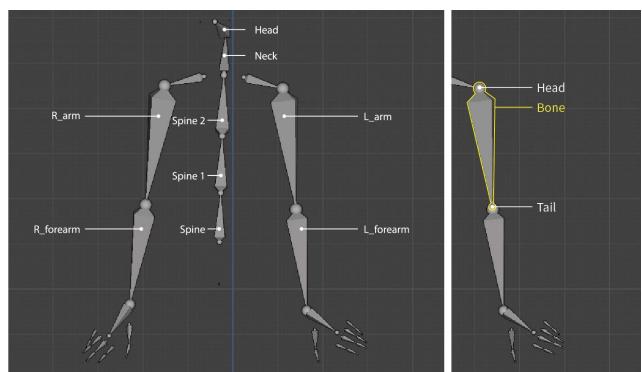


FIGURE 6. The structure of motion capture data.

G. DERIVATION OF GESTURE DATA AND APPLICATION OF METHODS

Our objective was to derive motion data based on the three aforementioned methods. For instance, the principal gesture of the “loving” emotion, designated as Pose 1, is a waving motion with the right hand raised, occurring between Frames 1 and 44. In accordance with the application method for joint data matching, we analyzed the R_arm and R_forearm joints from the motion capture data and applied the results to the robot’s right Joint B.

1) DERIVING THE ORIENTATION OF KEY POSE 1 FOR THE “LOVING” EMOTION

A detailed examination of the motion patterns exhibited by R_arm and R_forearm revealed that both bones demonstrated the greatest degree of movement along the x-axis. Therefore, we concluded that the principal motion direction of Key Pose 1 for the “loving” emotion was the x-axis. Consequently, we deemed it appropriate to implement the motion along the x-axis when applying it to the robot gesture. In this instance, the x-axis motion of the right arm is to be understood as the pitch motion of the robot’s right Joint B.

2) DERIVING A RANGE FOR KEY POSE 1 OF THE “LOVING” EMOTION

To ascertain the range of motion, we first needed to determine the position coordinates of the head and the tail in both

Frame 1 and Frame 44 for each joint, specifically R_arm and R_forearm. The requisite position coordinates were extracted using the Python Script feature of Blender. Table 13 illustrates the position coordinates of the head and the tail in relation to the R_arm bone in both Frame 1 and Frame 44.

TABLE 13. R_arm bone coordinates in frame 1 and frame 44 unit: mm.

	Position Coordinates of R_arm in Frame 1			Position Coordinates of R_arm in Frame 44			Unit: mm
	X	Y	Z	X	Y	Z	
Head	-12.8831	5.1044	144.8714	-10.0280	6.9217	145.1778	
Tail	-31.0556	9.2815	125.4850	-21.7089	-12.8643	131.1925	

Using the coordinates obtained, we constructed a 3D representation of the two lines: Line 1 connecting the head and tail coordinates of Frame 1 and Line 2 connecting the head and tail coordinates of Frame 44.

To calculate the angle of movement from Line 1 to Line 2, the concept of vectors was utilized since vectors are defined solely by direction and magnitude. First, Line 1 and Line 2 were plotted in a 3D space. Then, Line 2 was aligned to intersect with Line 1 while maintaining the same direction and magnitude. The angle of the intersecting line (Line 3) with Line 1 was then measured. As a result, the movement angle of the R_arm was determined to be 56.3 degrees.

Hence, the range of the right-hand raising motion in Key Pose 1 of R_arm was determined to be 56.3°. Similarly, the range of R_forearm, the other joint used in Key Pose 1, was calculated and found to be 158.2°. Fig. 7 provides a visual representation of two straight lines in 3D space for R_arm, along with the range between the two lines.

In accordance with the methodology of range factor analysis, the larger of the two angles R_arm and R_forearm—that is, 158.2° of R_forearm—was applied to the right arm of the robot, Joint B. However, since the maximum movement angle of the x-axis (pitch) for Joint B was 120°, the angle was adjusted to the robot's maximum rotation range of 120°.

3) DERIVING THE SPEED OF KEY POSE 1 FOR THE “LOVING” EMOTION

The speed calculation was based on the principles of speed factor analysis. The speed was calculated by dividing the amount of range change by the amount of frame change. The range previously determined was 120°, which could be

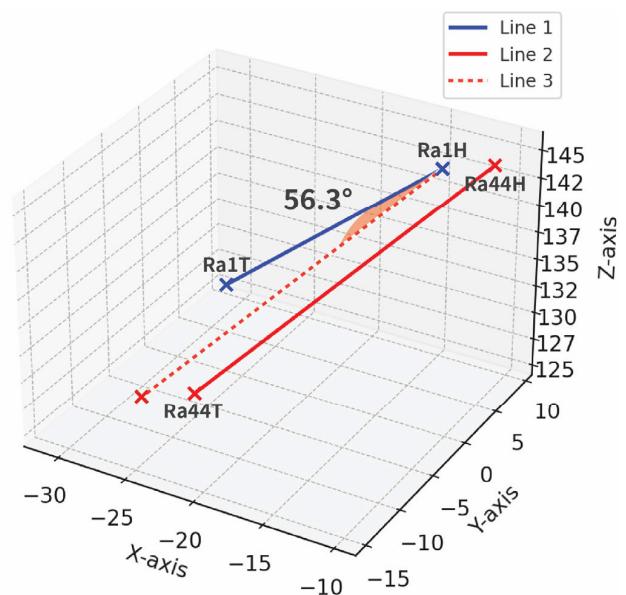


FIGURE 7. R_arm 3D lines visualization and angle. Ra1H represents the head coordinate of R_arm in Frame 1, while Ra1T represents its tail coordinate. Similarly, Ra44H represents the head coordinate of R_arm in Frame 44, and Ra44T represents its tail coordinate. Line 1 is the straight line connecting Ra1H and Ra1T, whereas Line 2 connects Ra44H and Ra44T. Line 3 is a straight line generated by translating Line 2 to intersect with Line 1, considering only the vector properties, such as direction and magnitude, while disregarding its position. Consequently, the angle between Line 1 and Line 3 was measured to be approximately 56.3°.

converted to approximately 2.09 radians. Accordingly, the speed was calculated as 2.09 divided by the ratio of 43 to 30, equaling 1.45 radians per second.

4) DERIVING THE ACCELERATION OF KEY POSE 1 FOR THE “LOVING” EMOTION

The application of acceleration was based on the Euler rotation graph, which is a visual representation of the angle of rotation about a particular axis over time. The direction of movement of Key Pose 1 was aligned with the x-axis, and the range of motion of the R_forearm bone was applied to the robot. Consequently, the X Euler rotation graph of R_forearm, as shown in Fig. 8, was derived, and acceleration was applied to the robot's gesture by reflecting the pattern of this graph in the motion Bézier curve.

5) APPLYING ADDITIONAL COMPENSATION TO KEY POSE 1 OF THE “LOVING” EMOTION

A hand wave with the right hand raised was the optimal gesture video and key pose description for the “loving” emotion. However, due to the inherent limitations of motion capture technology, the dataset was not detailed enough to fully capture the nuances of the hand-waving motion. Therefore, the hand-waving motion was further refined by referring to the optimal gesture video, contingent upon the manner in which the supplementary correction was implemented.

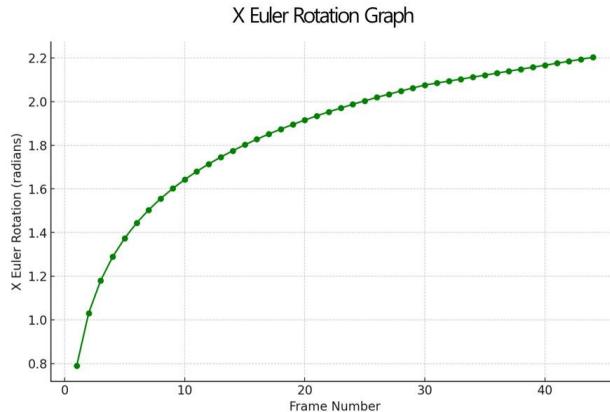


FIGURE 8. The graph of the X Euler rotation of R_forearm from frame 1 to frame 44. The graph illustrates the rotation angle of R_forearm along the x-axis as a function of the frame. It should be noted that there is a discrepancy between this graph and the range of motion derived earlier between the two straight lines in 3D space. Nevertheless, in this study, the objective of deriving acceleration was to facilitate the robot's ability to implement acceleration patterns analogous to those observed in humans, based on the pattern of the acceleration graph rather than the precise numerical application of acceleration. Accordingly, we focused on the shape and pattern of the curve over time, which we then applied to the robot's acceleration curve in a similar manner.

H. DERIVATION OF MOTION DATA FROM ROBOT EMOTIONS

1) DERIVING MOTION DATA FOR THE “LOVING” EMOTION
We analyzed the motion data for all key poses of the “loving” emotion, in accordance with the aforementioned methods, and derived the range, speed, direction, and acceleration of the “loving” emotion gesture. These values were applied to the robot by applying each method. Table 14 represents the joints, directions, ranges, and speeds of each key pose applied to the robot, while Fig. 9 illustrates the acceleration in a graphical format.

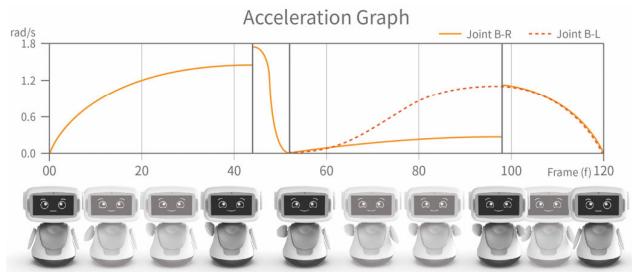


FIGURE 9. Acceleration graph of the “loving” emotion. We derived accelerations based on patterns in the euler rotation graph and applied them to the robot body acceleration graph.

2) DERIVING GESTURE DATA FOR THE REMAINING EMOTIONS

Similarly, we obtained gesture data for the emotions “joyful,” “upbeat,” “hopeful,” “concerned,” and “grateful,” following the previously described methodology, as summarized in Table 15.

TABLE 14. The range, direction, and speed for the “loving” emotion.

Key Pose (Frame)	Robot Joint Movement	Robot Image
Key Pose 1 (1-44) Raise your right hand and wave	Joint B (R): Pitch Range +120° Speed 1.45 rad/s	
	Joint B (R): Roll Range +/- 15° Speed 1.45 rad/s	
Key Pose 2 (45-52) Lower your right hand slightly	Joint B (R): Pitch Range -23.6° Speed 1.75 rad/s	
Key Pose 3 (53-98) Raise both hands and wave	Joint B (R): Pitch Range +12.6° Speed 0.14 rad/s	
	Joint B (L): Pitch Range +95.2° Speed 1.10 rad/s	
Key Pose 4 (99-120) Lower both hands	Joint B (R/L): Roll Range +/- 15° Speed 0.14 rad/s	
	Joint B (R): Pitch Range -109° Speed 1.10 rad/s	
	Joint B (L): Pitch Range -95.2° Speed 1.10 rad/s	

VII. VALIDATION EVALUATION OF THE FACIAL EXPRESSION AND GESTURE DESIGN OF A SOCIAL ROBOT FOR THE ELDERLY

A. PREFERENCE EVALUATION OF INTEGRATED FACIAL EXPRESSION AND GESTURE DESIGN

We conducted a user preference evaluation to ascertain whether the facial expressions and gestures of our social robot for seniors are appropriate for interacting with users and can effectively convey emotions. A total of 25 potential elderly users were presented with a video of the HRI design, which depicted the robot’s facial expressions and gestures. Subsequently, they rated the video on a seven-point scale and provided their opinions regarding potential areas for improvement. The evaluation results are shown in Table 16. The P-value was less than 0.01, confirming that the preference evaluation results were statistically significant. Facial expressions and gestures that received low scores were adjusted based on user feedback. Specifically, the size, shape, and angle of facial components were modified, while gestures were adjusted in terms of speed, duration, and range of motion.

B. MODIFICATION OF INTEGRATED FACIAL EXPRESSION AND GESTURE DESIGN

The “joyful” emotion facial expressions and gesture design received a low average rating of 4.5 on a seven-point scale. The users indicated that the facial expressions were overly

TABLE 15. Direction, range, and speed data for the remaining emotions.

Robot Emotions	Key Pose (Frame)	Robot Joint Movement	
Joyful	Key Pose 1 (1-67) Raise your arms high above your head	Joint B (R/L): Pitch	
		Range	+120°
		Speed	0.95 rad/s
	Key Pose 2 (68-103) Lower your arms	Joint B (R/L): Pitch	
		Range	-147°
		Speed	2.19 rad/s
		Joint B (R/L): Roll	
		Range	+/- 30°
	Key Pose 3 (104-136) Stop the motion	No movement	
		Joint C: Yaw	
Upbeat	Key Pose 4 (137-191) One turn to the right while raising and lowering your arms slightly.	Range	+360°
		Speed	3.48 rad/s
		Joint B (R/L): Roll	
		Range	+/- 40°
		Speed	0.98 rad/s > 2.07 rad/s
		Joint B (R/L): Pitch	
		Range	+27°
	Key Pose 1 (1-61) Raise both hands	Joint B (R/L): Pitch	
		Range	+120°
		Speed	1.02 rad/s
Hopeful	Key Pose 2 (62-91) Nod your head twice while keeping your hands still	Joint A: Pitch	
		Range	+/- 10°
		Speed	0.72 rad/s
	Key Pose 3 (92-122) Lower your hands	Joint B (R/L): Pitch	
		Range	-120°
		Speed	2.09 rad/s
	Key Pose 1 (1-33) Raise both arms	Joint B (R/L): Pitch	
		Range	+90.2°
		Speed	1.47 rad/s
		Joint B (R/L): Pitch	
		Range	+/- 15.3°
		Speed	1.56 rad/s
		Joint B (R/L): Pitch	
Concerned	Key Pose 3 (58-89) Lower your arms	Range	-73.3°
		Speed	1.22 rad/s
		Joint B (R/L): Pitch	
	Key Pose 4 (90-112) Raise your arms again	Range	+87.2°
		Speed	2.07 rad/s
		Joint B (R/L): Pitch	
	Key Pose 5 (113-123) Keep your arms up and sway slightly	Range	+/- 15.2°
		Speed	0.75 rad/s
		Joint B (R/L): Pitch	
	Key Pose 6 (124-141) Lower your arms	Range	-104.1°
		Speed	3.07 rad/s
Grateful	Key Pose 1 (1-27) Nod	Joint A: Pitch	
		Range	+/- 10°
		Speed	0.36 rad/s > 0.46 rad/s
	Key Pose 2 (28-84) Raise your arms up	Joint B (R/L): Pitch	
		Range	+120°
		Speed	1.11 rad/s
	Key Pose 3 (85-125) Stop the motion	No movement	
		Joint C: Yaw	
		Range	+/- 13.5°
	Key Pose 4 (126-144) Turn your head slightly to the right and then return it to its original position	Speed	0.76 rad/s > 0.86 rad/s
		Joint B (R/L): Pitch	
		Range	-120°
	Key Pose 5 (145-164) Lower your arms	Speed	0.79 rad/s
		Joint B (R): Pitch	
		Range	+28.7°
Grateful	Key Pose 1 (1-35) Raise your right arm	Speed	0.42 rad/s
		Joint B (L): Pitch	
		Range	+35.5°
	Key Pose 2 (36-64) Raise your left arm	Speed	0.65 rad/s

TABLE 15. (Continued.) Direction, range, and speed data for the remaining emotions.

	Key Pose 3 (65-74) Stop the motion	No movement	
	Key Pose 4 (75-109) Tilt your head slightly down while keeping your arms still	Joint A: Pitch	
	Range	-10°	
	Speed	0.15 rad/s	
	Key Pose 5 (110-128) Raise your head	Joint A: Pitch	
	Range	+10°	
	Speed	0.28 rad/s	
	Key Pose 6 (129-141) Tilt your head to the left	Joint C: Yaw	
	Range	+16.7°	
	Speed	0.72 rad/s	
Key Pose 7 (142-160) Lower both arms	Joint C: Yaw		
	Range	-16.7°	
	Speed	0.48 rad/s	
	Joint B (R): Pitch		
	Range	-28.7°	
Key Pose 7 (142-160) Lower both arms	Speed	0.83 rad/s	
	Joint B (L): Pitch		
	Range	-35.5°	
	Speed	1.01 rad/s	

TABLE 16. User preference evaluation results.

Robot Emotions	Preference Rating Score Total	Average Score
Loving	147	5.9
Joyful	112	4.5
Upbeat	144	5.8
Hopeful	146	5.8
Concerned	117	4.7
Grateful	102	4.1

exaggerated compared to the gestures. Additionally, they noted that the overall speed of the gestures was insufficient to convey the intended emotion. In response, we slightly adjusted the squint of the eyes and the size of the mouth in the facial expression design. For the gesture design, we modified the speed of Key Pose 1 and Key Pose 2 to be slightly faster than the original speed. In addition, we adjusted the overall speed by shortening the duration of the pause in Key Pose 3.

The “concerned” emotion facial expressions and gesture design received a low average rating of 4.7. The users indicated that the facial expressions, particularly the angles of the eyebrows and the shapes of the eyes and mouth, were overly exaggerated, leading to an imbalance with the gestures. Additionally, they noted that the gestures lacked sufficient expression of concern or anxiety. In response to these findings, we made slight adjustments to the angles of the eyebrows and smoothed their curves in the facial expression design. The curve of the mouth was softened, and the shape of the eyes was refined. For the gesture design, we modified the number of gesture repetitions and the duration of the pause, with the aim of enabling the robot to more accurately convey states of anxiety and concern. The duration of the pause in Key Pose 3 was reduced by over 50% to minimize static

motion. Furthermore, the rotation in Key Pose 4 was modified to be repeated twice.

The “grateful” emotion facial expressions and gesture design received a low average rating of 4.1. The users indicated that the facial expressions were overly intense compared to the gestures. Additionally, the arm movements were deemed to be inconspicuous, and the rotation was perceived as unnatural. In response to these observations, we adjusted the position of the eyebrows and refined the curve of the eyes to create more natural facial expressions. For the gesture design, we nearly doubled the range of the arm movements and removed the unnatural rotation to make the emotion seem more natural.

C. REEVALUATION OF INTEGRATED FACIAL EXPRESSION AND GESTURE DESIGN

The revised “joyful,” “concerned,” and “grateful” facial expressions and gesture designs were subjected to a second evaluation by a sample of 25 prospective users of the social robot. Following the presentation of the revised gesture designs, the participants indicated their preferences on a seven-point scale. For all designs, the average score was 5.5 or higher, as shown in Table 17. The evaluation results had a P-value that was less than 0.05, indicating that the preference results were statistically significant.

TABLE 17. Reevaluation results of the preferences for the modified design.

Robot Emotions	Reassessment Score Total	Average Score
Joyful	139	5.6
Concerned	155	6.2
Grateful	146	5.8

The social robot facial expression and gesture designs for the six emotions derived in this study were found to be positively evaluated by prospective users. The finalized designs can be seen in the following video: <https://youtu.be/luEXE0CJ6NU>

VIII. CONCLUSION

The objective of this study was to devise facial expressions and gestures for a social robot intended for use by the elderly, with the aim of enabling natural and humanlike interactions through the accurate representation of the robot’s emotions.

In this study, we first organized the principal elements of HRI design into four categories: gestures, facial expressions, appearance, and sound. Subsequently, we derived a detailed HRI design centered on facial expressions and gestures that are unique to robots.

We have analyzed human facial expressions and gestures and applied them to robots. Our objective is to provide guidelines for the design of natural, human-friendly facial expressions and gestures for social robots and thus facilitate deep emotional connections between these robots and elderly users. This can enable the transition of robots from mere tools

to companions that inspire trust and familiarity. In particular, the expression of emotions through a combination of facial expressions and gestures has the potential to enhance communication between robots and users, thereby fostering emotional connection. Our findings can inform improvements to the HRI design of social robots, enhancing their functionality and facilitating their transition from the research and development phase to commercialization.

Nevertheless, this study is constrained by the fact that it presents a single facial expression and gesture for each emotion, which does not account for the potential variations in emotional intensity. Future research should develop a series of facial expressions and gestures that will categorize emotional intensity and further refine the emotional expression of social robots. We argue that this will establish a new standard for HRI.

Additionally, the facial expressions and gestures proposed in this study were designed based on the preferences of a specific user group. However, emotional expressions can vary in interpretation and application depending on cultural differences and individual characteristics. In this context, the process outlined in this study can be adapted to accommodate the unique traits and cultural backgrounds of individual users, thereby expanding the applicability of HRI design across diverse social contexts.

REFERENCES

- [1] C. Kim, “Behavior expression technology of social robots,” *J. Korea Robot. Soc.*, vol. 14, no. 4, pp. 25–36, Oct. 2017.
- [2] L. Li, Y. Li, B. Song, Z. Shi, and C. Wang, “How human-like behavior of service robot affects social distance: A mediation model and cross-cultural comparison,” *Behav. Sci.*, vol. 12, no. 7, p. 205, Jun. 2022, doi: [10.3390/bs12070205](https://doi.org/10.3390/bs12070205).
- [3] L. F. Barrett, “Are emotions natural kinds?” *Perspect. Psychol. Sci.*, vol. 1, no. 1, pp. 28–58, Mar. 2006.
- [4] L. A. Camras, “Differentiation, dynamical integration and functional emotional development,” *Emotion Rev.*, vol. 3, no. 2, pp. 138–146, Apr. 2011.
- [5] S. Hong, *Robot Design Practice*. Siheung, South Korea: Korea Polytech. Univ. Press, 2022, p. 74.
- [6] J. H. Lee and H. R. Kim, “A study on the design of home social robots reflecting nonverbal interaction characteristics,” *J. Basic Des. Art*, vol. 24, no. 1, pp. 243–258, Jan. 2023.
- [7] H. S. Soo, S. C. Heo, E. Kim, and Y. J. Chang, “A study on the practical human–robot interface design for the development of shopping service support robot,” *Arch. Des. Res.*, vol. 19, no. 4, pp. 81–90, Apr. 2006.
- [8] O.-J. Kwon, Y.-M. Lee, D.-S. Kim, O.-K. Lee, and J.-S. Yim, “Developing scenario for elderly residents’ behaviors at home using persona-based scenario method,” *J. Korean Housing Assoc.*, vol. 28, no. 3, pp. 65–74, Jun. 2017, doi: [10.6107/jkha.2017.28.3.065](https://doi.org/10.6107/jkha.2017.28.3.065).
- [9] M. A. Brackett, *Permission to Feel*. New York, NY, USA: Celadon Books, 2019.
- [10] Paul. (Apr. 2024). *Comparison of DALL-E 3 and Midjourney: Analyzing AI Image Generation Tools*. Paul’s Media Set. [Online]. Available: <https://paulsmmediaset.com>
- [11] A. J. Yap. (Oct. 2023). *DALL-E 3 vs. MidJourney: A Side-by-Side Quality Comparison*. Gold Penguin. [Online]. Available: <https://goldpenguin.org/blog/dalle3-vs-midjourney/>
- [12] P. Ekman and W. V. Friesen, *Unmasking the Face: A Guide to Recognizing Emotions From Facial Expressions*. Los Altos, CA, USA: Malor Books, 2003.
- [13] C. Ilyas, R. Nunes, K. Nasrollahi, M. Rehm, and T. Moeslund, “Deep emotion recognition through upper body movements and facial expression,” in *Proc. 16th Int. Joint Conf. Comput. Vis., Imag. Comput. Graph. Theory Appl.*, 2021, pp. 669–679.

- [14] H. S. Jeong and S. S. Hong, "A study on motion HRI design of educational robot through user preference analysis," *J. Des. Converg.*, vol. 21, no. 5, pp. 1–16, Oct. 2022.
- [15] H. S. Jung and S. S. Hong, "A study on the development of optimal motion HRI design for café service robots," *J. Korean Soc. Des. Sci.*, vol. 37, no. 2, pp. 187–207, Apr. 2024.
- [16] J. S. Song, C. W. Kim, and H. S. Yang, "Humanoid's imitation of human motion by using motion capture," in *Proc. Korean Inst. Electr. Engineers Conf.*, Jun. 2006, pp. 1939–1940.
- [17] S. S. Kim, C. W. Kim, J. H. Park, and B. J. Yoo, "Upper body motion reproduction of humanoid robots using motion capture database," in *Proc. Korean Inst. Electr. Engineers Conf.*, Jul. 2006, pp. 1935–1936.
- [18] K. Yamane, Y. Ariki, and J. K. Hodgins, "Animating non-humanoid characters with human motion data," in *Proc. ACM SIGGRAPH/Eurographics Symp. Comput. Animation (SCA)*. Madrid, Spain: Eurographics Assoc., Jul. 2010, pp. 169–178.
- [19] H. Knight and R. Simmons, "Laban head-motions convey robot state: A call for robot body language," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2016, pp. 2881–2888.
- [20] R. von Laban, *Modern Educational Dance*, 3rd ed., London, U.K.: Macdonald & Evans, 1975.
- [21] A.-A. Samadani, S. Burton, R. Gorbet, and D. Kulic, "Laban effort and shape analysis of affective hand and arm movements," in *Proc. Humaine Assoc. Conf. Affect. Comput. Intell. Interact.*, Sep. 2013, pp. 343–348.



SEONGSOO HONG received the Ph.D. degree in industrial design from Hanyang University, Seoul, South Korea, in 2004.

Since 2004, he has been a Professor with the Department of Design Engineering, Tech University of Korea, Siheung, South Korea. He has been leading the Robotics Design Laboratory focused on industrial design, since 2005. He has a large portfolio of robot design projects, including the Yeosu Expo Humanoid Robot (NAVI), industrial robots, underwater exploration robots, patrol robots, entertainment robots, and home service robots. His research interests include product design, robot design, HRI design, and industrial equipment.

• • •



SUJIN JO received the B.S. degree in industrial design engineering from the Tech University of Korea, Siheung, South Korea, in 2023, where she is currently pursuing the M.S. degree in industrial design engineering. Her research interests include robot design, human–robot interaction (HRI) design, and product design.