

Emoji-Inspired Continuous Emotional Expression for Aerial Robots Using Eye Displays Based on the VAD Model

Ayano Miyamichi¹ and Kei Okada¹

Abstract—In recent years, emotion expression systems have attracted growing attention in the field of Human-Drone Interaction. However, most of these systems adopt discrete expression models, which are insufficient for capturing the ambiguity and transience of emotions. On the other hand, approaches based on continuous emotion models often rely on realistic and high-degree-of-freedom facial expressions, which are difficult to implement on non-humanoid and weight-constrained platforms such as aerial robots. To address this challenge, we propose a visual interface that continuously expresses emotion only through the movement of the eyes. In designing eyes expressions, we consider the non-biological appearance of aerial robots and emphasize friendliness and universality through an abstract visual style, inspired by the visual characteristics of emoji. Furthermore, we adopt Valence-Arousal-Dominance(VAD) model, a continuous emotional space comprising three psychological dimensions to guide the expressive behavior. Based on this model, we designed and implemented an eye display system that integratively controls visual components in accordance with the emotional state. Finally, we installed this system on a real robot, and verified its effectiveness in emotion recognition and the appropriateness of its expression through evaluation experiments. The results confirmed its effectiveness, particularly in conveying high-arousal emotional states.

I. INTRODUCTION

Aerial robots have the advantage of being able to move freely in disaster sites and crowded spaces where it is difficult for humans and ground robots to enter due to their high maneuverability [1], [2]. They are widely used in non-contact tasks such as aerial imaging [3], surveillance [4], inspection [5] and search operations [6]. In recent years, the use of aerial robot mobility for daily human support has gained attention, giving rise to Human-Drone Interaction (HDI) research aimed at natural interaction [7]–[10]. However, in line with their typical applications, most of the HDI studies have focused on non-contact scenarios, such as flight-based interaction [7]–[9] or limited touch-based approach [10], and the potential of utilizing aerial robots' mobility in close proximity to humans remains limited.

To address this, we have focused on perching, which means a bird-inspired behavior of landing and grasping, and has proposed a new framework in which aerial robots initiate interaction by landing on the user's arm or body. This physical contact enables close-proximity interaction, allowing for bidirectional exchange of information, such as sensing the user's physical or emotional state and providing

*This work was not supported by any organization

¹The authors are with the Department of Mechano-Infomatics, Graduate School of Information Science and Technology, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, 113-8656, Japan. miyamichi@jsk.imi.i.u-tokyo.ac.jp

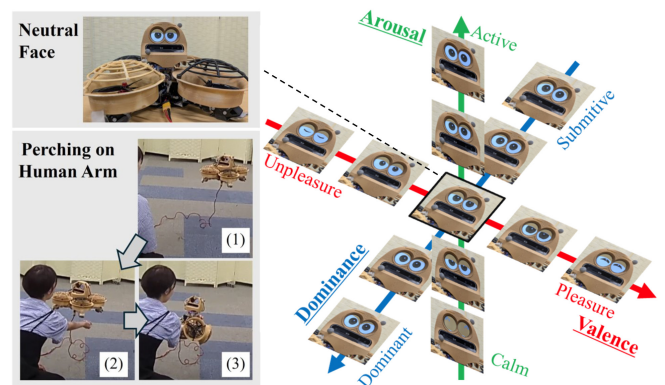


Fig. 1. Perching aerial robot with continuous eye expressions in the VAD space

visual or haptic feedback. In such close-range and embodied interactions, it is crucial for aerial robots to express emotions in order to clearly communicate their status and intentions to provide assistance. For example, by changing its facial expression during perching, the robot can indicate the timing of perching, helping the user anticipate contact. Moreover, expressing emotions during approach and perching phases can relieve anxiety and foster a more intimate and trustful relationship between the user and the robot. However, their non-biological appearance and strict weight limits make expressive communication difficult, highlighting the need for lightweight and intuitive visual cues. Therefore, this study aims to realize emotional expressions that are intuitive and capable of continuous change, even with lightweight, non-biological forms, by introducing a visual expression method using "eyes". Specifically, drawing inspiration from emoji as a culturally embedded visual language, we map them onto the Valence-Arousal-Dominance (VAD) emotional model [11] and translate these values into eye movements and visual effects to construct a visual expression model that enables continuous and smooth emotional expression. Furthermore, we verify the effectiveness of the proposed model through experiments. This enables aerial robots to convey internal states through intuitive, nonverbal cues, facilitating smooth and socially acceptable close-range interaction.

II. RELATED WORK

A. Emotional Expression in Social Robots

In recent years, social robots [12], which are designed for social and emotional interaction, have attracted attention in the field of Human-Robot Interaction (HRI). Social robots

are already being used in medical and educational institutions, public facilities, and homes [13], where appropriate emotional expression is essential for natural communication. Eyssel et al. [14] showed that robots expressing emotions are more easily anthropomorphized and their intentions better understood, highlighting the importance of emotional expression in effective HRI.

Traditionally, most research on emotional expression in social robots has been based on Ekman's six basic emotion (joy, sadness, anger, fear, surprise, and disgust) [15] using a discrete emotion model [16]. However, several studies have noted that this model may not fully capture the complexity and subtleties of human emotional experiences [17]. To address this limitation, some researchers have adopted dimensional models such as Russell's Circumplex Model, which represents emotions using two dimensions: Valence (positive-negative) and Arousal(active-calm) [18]. Mehrabian model, further introduces a third dimension, Dominance, to describe emotions in a three-dimensional space [11]. These frameworks enable smooth and continuous emotional representation. For example, Kismet [19] expresses emotions along valence, arousal, and stance dimensions, enabling fluid transitions via spatial interpolation. Similarly, Han et al. [20] mapped robot emotions onto a Valence-Arousal space and generated facial animations based on mood and personality. While such continuous models support subtle and transitional expressions, their implementation often requires complex hardware, such as movable facial components, that is unsuitable for aerial robots where light weight is critical. Moreover, realistic face designs risk evoking the uncanny valley effect [21], making them less appropriate for aerial platforms with non-humanlike appearances.

B. Emotional Expression in Human-Drone Interaction

Human-Drone Interaction (HDI), following the definition of HRI, focuses on design and evaluation of aerial robot systems intended for collaboration with human users [22]. In particular, research on social drones for everyday environments has been advancing, exploring how aerial robots, whose physical forms and motion characteristics differ from human or animal, can effectively convey emotions.

One approach utilizes flight behavior as a form of bodily expression. For example, Arroyo et al. [23] used combinations of head movement, eye color, and propeller sound to convey positive, neutral, and negative emotions. Cauchard et al. [24] showed that three personality traits could be expressed through variations in flight path and speed, and evaluated their recognizability. Hieida et al. [25] proposed a method to express emotions based on Japanese idioms for joy, anger, sadness, and pleasure, using dynamic parameters such as velocity, acceleration, and displacement. These studies are notable in that they use flight behavior as bodily expression. However, the range of emotions expressible in this manner remains discrete and limited.

Another approach incorporates facial elements into aerial robots to support emotional expression. For instance, Ruijten et al. [26] found that even simple, non-functional "eyes"

enhanced perceived friendliness. Herdel et al. [27] implemented a cartoon-style face and presented Ekman's six basic emotions using combinations of eyes, eyebrows, pupils, and mouth with three intensity levels. They demonstrated that static images conveyed five emotions (joy, sadness, fear, anger, surprise) and videos conveyed four (joy, surprise, sadness, anger) with high recognition accuracy. Furthermore, Bretin et al. [28] investigated how digital facial expressions based on Herdel et al.'s work [27] and gaze, influenced spatial perception in VR, showing that a aerial robot's gaze and facial expressions significantly affect human psychological receptivity. While these studies highlight the value of visual expressions as social signals, most rely on discrete emotion models. In fact, 76% of studies on emotional expression in autonomous systems, including aerial robots, adopted Ekman's six basic emotions [29]. Moreover, much of this work has been conducted in simulated environments, with limited implementation on physical aerial robots operating under real-world constraints. To enable more natural and flexible emotional communication, there is a growing need for hardware-integrated approaches that express emotions continuously and in coordination with real-world behavior.

C. Positioning and Contribution

As mentioned above, many studies on emotional expression in social robots have relied on discrete models, which often struggle to capture the ambiguity and transitional nature of emotions. Similarly, facial expression systems for aerial platforms have also remained largely discrete and limited in real-world integration. In contrast, continuous models offer greater flexibility but typically require complex mechanical structures that are unsuitable for lightweight aerial robots.

This study targets a perching aerial robot capable of landing on and grasping a user's arm or body to enable close physical interaction. Such robots must maintain a lightweight structure to preserve flight and perching functionality, while also enabling expressive emotional communication. To address this, we propose a lightweight visual interface based on eye movements as an intuitive means of emotional expression, tailored to aerial robots. Considering their non-biological appearance, we adopt an abstract facial design that emphasizes friendliness and universality. The visual design is inspired by the expressive features of emoji, which offer a simplified yet effective means of conveying emotion.

For emotional modeling, we employ the continuous VAD (Valence-Arousal-Dominance) model [11] and map facial-style emoji onto this space to design eye components. Based on this mapping, we implement a control system that continuously adjusts iris size, eyelid position, and visual effects according to emotional state, and integrate it into a physical aerial robot to assess the recognizability and effectiveness of the expressed emotions.

This study makes the following contributions:

- We propose a new approach to robot emotion expression by mapping emoji semantics onto the continuous VAD emotional space, providing a psychologically grounded basis for smooth and continuous visual expression.

- We design and implement a lightweight eye-display system that realizes continuous emotional transitions by mapping VAD values to the control of eye elements.
- We demonstrate the first physical realization of a continuous VAD-based emotional expression system on an aerial robot and empirically confirm its effectiveness through user experiments, particularly in conveying high-arousal emotions.

In summary, this study presents the first implementation of a continuous emotional expression framework for aerial robots that connects psychological models, emoji-based visual abstraction, and lightweight hardware design, offering a physically feasible and engaging approach to HDI.

III. CONTINUOUS EYE-BASED EMOTIONAL EXPRESSION DESIGN USING AN EMOJI-DERIVED AFFECTIVE SPACE

A. Continuous Eye-Based Emotional Expression Design for Aerial Robots

1) Expressing Rich Emotions with a Simple Configuration:

To achieve rich human-robot interaction with non-biological aerial robots, it is essential to use visual expressions that clearly convey intentions and states. However, strict payload constraints limit the equipment that can be installed while maintaining mobility. Therefore, this study focuses on the eyes, which most effectively convey human emotions [30], [31]. Even with simple designs, eye expressions can intuitively indicate emotional direction and, through subtle shape changes and animations, achieve emotional richness suitable for compact and lightweight platforms.

Conventional models often switch eye and mouth shapes based on discrete emotion categories (e.g., joy, sadness, anger), limiting smooth transitions and complex intermediate states. To address this, we adopt the Valence-Arousal-Dominance (VAD) model [11], which represents emotions continuously along three dimensions, Valence (pleasantness-unpleasantness), Arousal (active-calm), and Dominance (dominant-submissive). Applying this model enables continuous control of eye elements such as pupil size and eyelid motion, allowing natural and gradual emotional expressions even under hardware constraints.

2) *Expression Design Inspired by Emoji:* In this study, we develop a robot emotion model that applies the visual characteristics of emoji to the design of expressive interfaces. Emoji are widely used in modern digital communication, and face-style emoji convey relatively consistent meanings across cultures [32]. Owing to the stable correspondence between their visual features and emotional interpretations, emoji serve as effective visual media for emotional communication and have been actively studied in affective research, where some studies have mapped emoji into affective spaces for quantitative analysis [33]. Fischer et al. [34] further reported that emoji and human faces were rated higher in emotional intensity than emoticons, and that emoji were recognized as quickly as, or even faster than, human faces in most emotion categories except anger and fear. These findings indicate that emoji combine high emotional clarity with cognitive

efficiency, making them promising visual symbols for robot emotion expression. However, most previous studies have treated emoji as supplementary to text analysis, and few have applied them to robot visual expression. Among these, Gomez et al. [35] proposed an interface called "robomoji" for the telepresence robot Haru, which transforms emoji sent from a smartphone into visual and behavioral expressions on the robot. Yet this approach limited to the correspondence between discrete emojis and individual expressions.

From this perspective, this study focuses on emojis, which are culturally learned visual symbols, rather than realistic faces or character-like expressions, and attempts to design intuitive and universal emotional expressions for aerial robots with non-biological appearances.

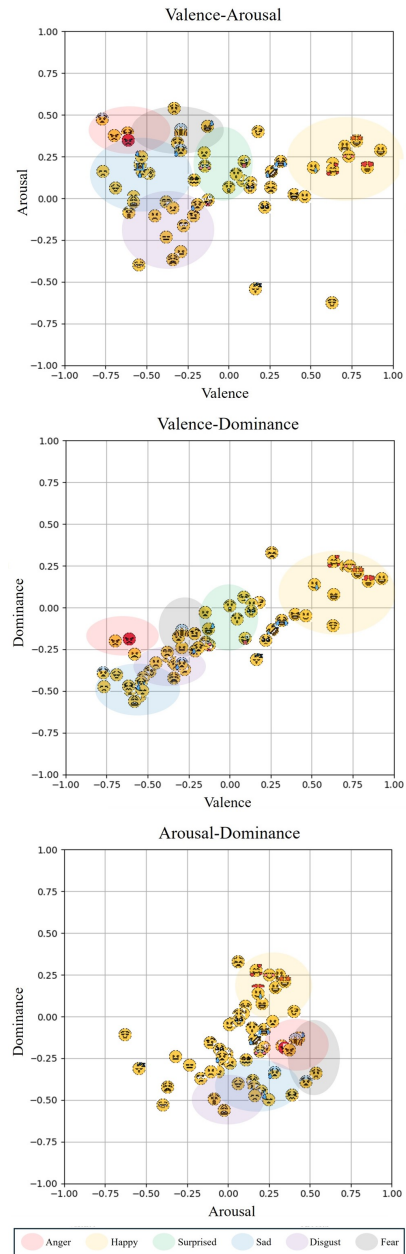


Fig. 2. Distribution of emoji in Valence-Arousal-Dominance (VAD) space

B. Design of a Visual Expression Model for Robot Based on Affective Space

To enable continuous emotional expression through the eyes of an aerial robot, this study focuses on emoji and constructs a three-dimensional affective space model based on their semantic structure. From this model, we derive design parameters for the robot's eye expressions.

1) Construction of an Emoji-Based VAD Affective Space:

First, we select 56 facial-expression emoji registered in the Unicode standard. Emoji with non-facial semantic elements—such as hand gestures (e.g., U+1F917), head effects (e.g., U+1F92F), or special components around the mouth (e.g., U+1F910) are excluded in order to focus solely on facial expressions. For semantic information, we refer to the English annotation data provided in the CLDR (Common Locale Data Repository) [36] by the Unicode Consortium. From each emoji's description, we extract emotion-related keywords while removing non-emotional terms (e.g., "face", "eye", "143"). For example, the emoji (U+1F600) is annotated with terms like cheerful, grinning, happy, laugh, and smile, from which only the emotional descriptors are retained.

Next, we assign VAD scores to each keyword using the NRC-VAD Lexicon, which provides Valence, Arousal, and Dominance values for over 20,000 English words and is widely used as a standard affective rating resource [37]. For emoji with multiple associated keywords, we compute the average VAD score across the keywords to define a single VAD vector per emoji. These vectors were then plotted in a three-dimensional VAD space as shown in Fig. 2. Through this process, the semantic nuances of each emoji were quantitatively mapped as continuous values, providing a foundation for generating facial expressions with smooth emotional transitions.

C. Design of an Emotion Expression Model Using Eyes

1) Hardware Configuration:

In this study, we design and implement a dual eye-display system as a means of visual emotional expression for aerial robots. Each eye module, consisting of a 15g circular display integrated with an ESP32 microcontroller, and together the two modules account for only 2.4% of the robot's total weight, demonstrating the system's lightweight design. The display adopts a layered rendering structure in which multiple image layers are superimposed to construct various eye expressions. As shown in Fig. 3, each expression is composed by overlaying the following image layers in order from the bottom:

- Eyeball (base) image: Background representing the white of the eye
- Iris image: Center of the eye, which can be scaled and translated to express emotional direction and gaze
- Highlight (reflection) image: Reflected light that gives the eye a sparkle and vitality
- Eye shape/effect images: Auxiliary visual expressions for emphasizing emotions such as anger and joy
- Eyelid image: Expresses emotional tone through opening/closing and angle changes

Each of these layers is stored as a sprite image in memory and can be independently transformed and re-ordered during rendering. This structure enables continuous motion and smooth transitions in the eye's visual expression.

The display is controlled by an ESP32 microcontroller, which is connected to the onboard PC of the aerial robot via I2C communication. The computed VAD values are published as ROS (Robot Operating System) topics and transmitted to the ESP32. The ESP32 then uses the received VAD values to dynamically control iris size and position, eyelid opening and blinking, and visual effects such as tears or hearts, updating the eye expression in real time.

The eye display is mounted as part of the robot's facial parts, whose expression also changes depending on the robot's flight attitude, as shown in Fig. 4. Specifically, the mouth area is visually designed such that when the robot tilts forward, it appears to be smiling, while tilting backward creates the appearance of a frowning mouth. This pose-dependent visual illusion enhances emotional expressiveness by leveraging the robot's movement in space.

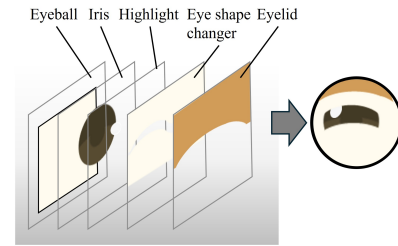


Fig. 3. Layered image composition for eye expression rendering

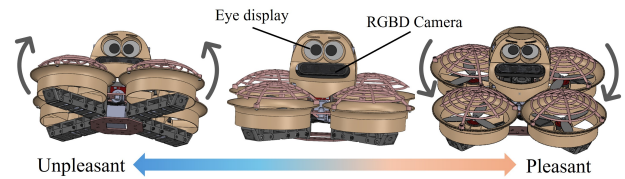


Fig. 4. Integration of an expressive eye display into a robot face that can change expression depending on the attitude.

2) Emotion Expression Control Based on VAD Values:

Based on the distribution of emoji in the VAD affective space, we extract characteristic tendencies along each dimension and design control parameters for the robot's eye expressions accordingly. We use three axes, Valence (V), Arousal (A), and Dominance (D), as continuous emotional values to dynamically change the iris size, eyelid shape, and expression effects. The control policies based on each axis are shown below.

• Valence (Pleasant-Unpleasant)

Valence is mapped to the size and shape of the iris. When the valence value (V) is low, the iris is rendered as a downward-convex shape; as V increases, the iris becomes larger, and when V exceeds a threshold, it is rendered as an upward-convex shape.

- $-0.5 \leq V \leq 0.5$: Adjusts the iris size in proportion to the value of V .
 $V = 0$ for standard size, $V > 0$ for enlargement, $V < 0$ for reduction.
- $V > 0.5$: Set the iris size to maximum and deform the eye shape to convex upward.
- $V < -0.5$: Set the iris size to minimum and deform the eye shape to convex downward.

- **Arousal (Active-Calm)**

Arousal is mapped to eyelid openness and visual effects. Higher arousal values (A) correspond to wider open eyes, while lower values result in a closed appearance. In addition, when A exceeds a certain threshold, symbolic emotional effects such as hearts or tears are added to visually enhance emotional intensity. A corresponds to the degree of eyelid opening (`upperlid_val`) and is determined by linear interpolation as follows:

- $A \leq -1.0$: Completely closed eyes (`upperlid_val` = 0.0)
- $A = 0$: Moderately open eyes (`upperlid_val` \approx 0.75)
- $A \geq 1.0$: Fully open eyes (`upperlid_val` = 1.0)

Furthermore, under certain conditions, a visual emotion enhancement effect (emoji-style decoration) is added:

- $V \geq 0.8$ and $0.4 < A < 0.6$: Heart
- $V \geq 0.8$ and $A \geq 0.6$: Star
- $-0.5 \leq V \leq -0.25$ and $0.0 \leq A \leq 0.5$ and $D < 0.0$: Tear
- $-0.5 \leq V \leq -0.25$ and $A > 0.5$ and $D < 0.0$: Fear
- $A \leq -0.9$: Sleepy

- **Dominance (Dominant-Submissive)**

Dominance is mapped to the horizontal offset of the eyelid, and asymmetrical tilting is used to add nuance to the expression:

- $D > 0$: The eyelid rises in a v shape, expressing anger or intimidation.
- $D < 0$: The eyelid descends in an inverted v shape, expressing sadness or a submissive impression.

By assigning each VAD dimension to distinct visual parameters, this system enables the robot to generate continuous, nuanced, and context-sensitive eye expressions, thereby achieving rich emotional expressiveness.

IV. EVALUATION OF VAD-BASED EMOTIONAL EXPRESSIONS FOR AERIAL ROBOTS

A. Experimental Validation of the Proposed Eye-Based Expression Model

To verify whether the expressions of aerial robots are appropriately recognized by humans, we conducted an evaluation experiment to verify the validity of emotional expressions in VAD space.

- **Experiment participants**
 Fourteen participants (engineering students and faculty members in their 20s and 30s) took part in the experiment on a voluntary basis. The participants included

undergraduate and graduate students as well as university faculty.

- **Experimental setup and procedure**

The aerial robot equipped with a face was placed on a table, and 19 designed expressions were presented continuously in a fixed sequence as shown in Fig. 5. Participants sat in groups of three or four people and evaluated the expressions while observing them from a position where the robot's expressions were clearly visible. Each participant received an individual questionnaire and was instructed to provide independent, subjective evaluations, without discussing or exchanging opinions with others. If desired, participants were allowed to request to revisit previous expressions or switch to a neutral face for confirmation.

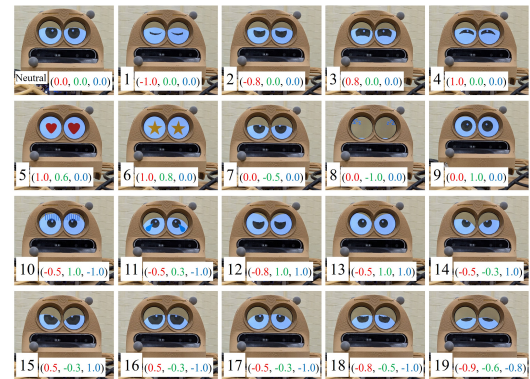


Fig. 5. The 19 expressions used in the experiment and their controlled VAD values

- **Questionnaire Structure**

The questionnaire consisted of two main questions and a free response section.

- **Question 1 (Selection of Emotional Words)**
 Based on the Japanese translation of Russell's circumplex model, participants were asked to select emotion words from a predefined list of 30 options (e.g., *aroused, surprised, excited, delighted, happy, funny, joy, glad, pride, neutral, tranquil, relax, calm, sleepy, lethargic, bored, depressed, lonely, sad, distress, upset, afraid, disgust, anxious, worried, angry, annoyed, tense, exhausted, miserable*) that they felt best matched each expression. If none of the listed words were appropriate, or if participants wished to elaborate on their impressions relative to other expressions, they were encouraged to provide additional comments in a free-response section.
- **Question 2 (Evaluation of the VAD three axes)**
 Participants were asked to evaluate each expression along the three axes of *Valence* (pleasant–unpleasant), *Arousal* (active–calm), and *Dominance* (dominant–submissive) using a 5-point Likert scale (1–5). Since some participants found the Dominance axis difficult to interpret, the following clar-

ification was provided in advance:

Dominance refers to the feeling of controlling one's environment, while passivity indicates a sense of being overwhelmed by others or the surroundings.

Finally, participants were asked to provide their overall impressions of the aerial robot's expressions and their thoughts on the changes in expressions in a free-response format.

B. Results

In this experiment, we presented expressions with varying control parameters and evaluated their impressions. The results are shown in Fig. 6 and Table I

First, for Expressions 1 to 4, we gradually changed the V value as the control value and changed the shape of the eyes. For Expressions 5 and 6, we added visual effects similar to those used in emojis by increasing the A value in addition to the control values used in Expression 4. The evaluation results showed that Expression 1, despite having a low V value and downward-curved eyes, was not perceived as unpleasant but rather gave the impression of "relax" or "sleepy" and the A values reported for low value. Furthermore, for Expression 2, while the intention was to convey "weavy face" based on emoji trends, it was actually perceived as other negative emotions such as "disgust".

On the other hand, for Expressions 3 to 6, most respondents accurately interpreted the intended emotions from the presented expressions, and reasonable trends were confirmed in the VAD values. In particular, the V values were high in the order of Expressions 3, 4, 6, and 5, which aligns with tendencies observed in emoji-based emotion mapping. In addition, regarding Expression 4, there were free comments such as "it was more favorable than Expression 3" and "it looked more pleasant than Expression 3". Regarding the A values, they were generally consistent except that Expression 5 had a higher value than Expression 6. Overall, Expressions 5 and 6 scored higher in A than 3 and 4, indicating successful control of arousal through visual enhancements.

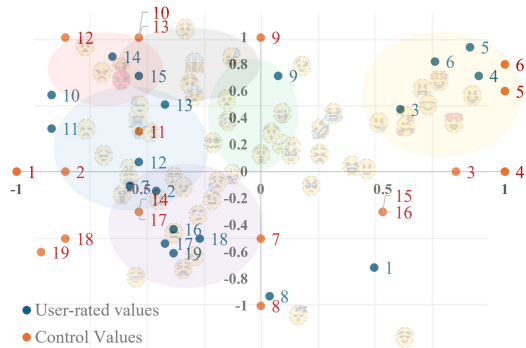


Fig. 6. Scatter plot comparing controlled and perceived Valence-Arousal values with background emoji distribution

Next, only the A values were changed for Expressions 7 to 9, and an effect was added to Expression 8. The questionnaire results showed that the intended emotions were ranked high, and the average A values were high in the order of 8, 7, and

	Intended Emotion	No.1	No.2
1	sad	relax(10)	sleepy(9)
2	annoyed, disgust	disgust(5)	sleepy, annoyed, bored, depressed(4)
3	happy (weaker than 4)	funny(8)	joy(6)
4	happy	funny(9)	happy, glad, delighted(8)
5	delighted	aroused, glad(10)	delighted(8)
6	excited	surprised, excited(8)	aroused(6)
7	sleepy (weaker than 8)	lethargic, bored(6)	disgust, sleepy(5)
8	sleepy	sleepy(14)	bored(7)
9	surprised	surprised, neutral(8)	tense, aroused(3)
10	afraid	worried(10)	afraid, upset(7)
11	sad	sad(14)	anxious, lonely(7)
12	angry	angry(6)	pride, bored, nervous, annoyed, aroused(2)
13	angry (weaker than 12)	angry, neutral, tranquil(3)	surprised, depressed, relax, worried, calm(2)
14	angry	annoyed(7)	angry(5)
15	angry	angry(6)	disgust(3)
16	sad	sleepy, depressed(5)	sad(4)
17	sad	lethargic(5)	sleepy, depressed, lonely(3)
18	disgust	sleepy(3)	lethargic, happy, calm, sad, anxious, worried, bored(2)
19	disgust	sleepy(6)	exhausted(4)

TABLE I
COMPARISON BETWEEN INTENDED EXPRESSIONS AND TOP TWO REPORTED EMOTIONS

9, indicating that the expression of emotions through arousal control was achieved. However, in Expression 7, the V value was low at -0.5 , so some respondents felt "disgust" and used expressions such as "scornful eye". In Expressions 10 and 11, the D value and V value were fixed, and effects were added by changing the A value. As a result, the emotions ranked at the top matched those intended, and in particular, Expression 10 elicited comments such as "shock" and "disappointment". The tear effect is considered to have effectively conveyed the emotion of sadness.

In Expressions 12 to 15, anger was expressed with a basic setting of low V values and high D values, and the strength of anger was shown in stages by changing the A value. Comparing Expressions 12 and 13, the difference in V values changed the shape of the eyes, and the survey results also showed that Expression 12 had a lower V value, indicating "stronger anger". Although the A value was fixed for control purposes, the survey results indicated that respondents felt that the shape of the eyes made them appear more alert. Furthermore, in Expression 14, lowering the A value slightly lowered the eyelids to partially cover the iris, and in Expression 15, increasing the V value enlarged the iris, resulting in higher overall VAD values than in Expressions 12 and 13. Responses indicated that Expression 14 was perceived as "stronger anger than Expressions 12 and 13", and Expression

15 was perceived as "stronger anger than Expression 14" or "pleasure".

Finally, Expressions 16 and 17 were designed to express sadness, while 18 and 19 targeted disgust. However, participant responses more frequently described these as "sleepy", "lethargic", "depressed" Interestingly, although these expressions were located in the third quadrant of the VAD space (low V , low A), they were more often labeled as "sleepy" than "sad". That said, "sad" was the second most frequent label for Expression 16, and responses to Expression 17 also included "sleepy", "depressed" and "lonely" suggesting a degree of successful emotional transmission despite some mismatch in categorical interpretation.

C. Discussion

This study demonstrated that visual expressions based on the VAD model enable aerial robots to convey a certain degree of diverse emotions using only their eyes. In particular, for expressions corresponding to emotions with high arousal values (anger, joy, excited, surprise, fear) and sleepiness, there was a high degree of consistency between the presented expressions and the subjects' perceptions, confirming the effectiveness of visual expressions. On the other hand, for emotions with low Valence values and low Arousal values, particularly expressions of "sadness" and "disgust", discrepancies were observed between the control values and the actual emotional cognition. For example, Expressions 1, 16, 17, 18, and 19 were designed to convey "sadness" or "disgust", but they were often interpreted "relaxed" or "sleepy". This may be due to the fact that the eye shapes appeared simply closed, which weakened the impression of negative emotion. Future improvements may include using downward-slanted eyelid shapes instead of convex forms to more clearly convey discomfort. In addition, adding a shape that wrinkles the brow between the eyebrows may also be effective in expressing disgust.

The addition of visual effects (e.g., tears, hearts, stars) proved highly effective in emphasizing emotional intensity. In particular, Expression 10, which included a tear effect, successfully conveyed emotions such as sadness and disappointment with high accuracy. This suggests that symbolic visual elements can complement eye shape to express emotions that are otherwise difficult to convey.

Regarding Dominance, comparisons between expressions with varied eyelid angles (e.g., Expressions 14 and 17) and iris sizes (e.g., Expressions 15 and 16) indicated that these parameters contributed to the perception of emotional intensity and assertiveness. In particular, iris enlargement had a positive effect on all Valence, Arousal, and Dominance values, making it an effective visual technique for enhancing the intensity and clarity of emotions. Furthermore, larger eyes tend to emphasize nuances such as "anger" in the angry expressions or "lack of confidence" in the sad expressions, and it was suggested that the the iris size may have an influence on Dominance rather than Valence.

As a future challenge, more effective visual strategies will be required to accurately convey low-Valence, low-Arousal

emotions such as "sadness" and "disgust". Possible improvements include raising the eyelids intentionally to avoid sleepy impressions, using inverted-V eyelid shapes to express sadness, and adding brow furrows to express disgust more clearly. Furthermore, we will explore adding tiny displays around the corner of mouth, controlled in coordination with the eye displays, to enhance subtle expressions and improve emotional clarity.

V. CONCLUSIONS

In this study, we propose a lightweight and intuitive method for emotional expression in aerial robots by mapping the affective structure of emoji onto the VAD model and applying it to eye-based visual expressions. By embedding emoji semantics in the VAD space, we derived visual parameters for continuous eye-based emotion display and verified their perceptual validity through user studies. Evaluation results confirmed that expressions and emphasis effects corresponding to high-arousal states were particularly effective, whereas discrepancies in recognition appeared in low-valence and low-arousal regions, indicating the need for further visual adjustments. These findings demonstrate that even aerial robots with non-biological appearances can achieve rich and empathetic emotional expression through emoji-based visual representation, providing new design guidelines for HDI.

Future work will extend the expressive range and evaluation of the system. We plan to integrate mouth expressions with the current eye-based design to improve recognition of low-valence, low-arousal emotions such as sadness and disgust. We also plan to expand user studies to include participants from broader demographic backgrounds and to conduct in-flight experiments with additional expressive cues such as sound and motion, evaluating performance under realistic HDI conditions.

REFERENCES

- [1] Dario Floreano and Robert J. Wood. Science, technology and the future of small autonomous drones. *Nature*, Vol. 521, pp. 460–466, 2015.
- [2] Vijay Kumar and Nathan Michael. Opportunities and challenges with autonomous micro aerial vehicles. *The International Journal of Robotics Research*, Vol. 31, No. 11, pp. 1279–1291, 2012.
- [3] Rogerio Bonatti, Yanfu Zhang, Sanjiban Choudhury, Wenshan Wang, and Sebastian Scherer. Autonomous drone cinematographer: Using artistic principles to create smooth, safe, occlusion-free trajectories for aerial filming. In *Proceedings of the 2018 international symposium on experimental robotics*, pp. 119–129. Springer, 2020.
- [4] Lefteris Doitsidis, Stephan Weiss, Alessandro Renzaglia, Markus W Ahtelik, Elias Kosmatopoulos, Roland Siegart, and Davide Scaramuzza. Optimal surveillance coverage for teams of micro aerial vehicles in gps-denied environments using onboard vision. *Autonomous Robots*, Vol. 33, pp. 173–188, 2012.
- [5] Francois Chataigner. Arsi: an aerial robot for sewer inspection. *Advances in Robotics Research: From Lab to Market*. Springer, Cham, 2020, pp. 249–274, 2020.
- [6] Nathan Michael, Shaojie Shen, Kartik Mohta, Yash Mulgaonkar, Vijay Kumar, Keiji Nagatani, Yoshito Okada, Seiga Kiribayashi, Kazuki Otake, Kazuya Yoshida, Kazunori Ohno, Eijiro Takeuchi, and Satoshi Tadokoro. Collaborative mapping of an earthquake-damaged building via ground and aerial robots. *Journal of Field Robotics*, Vol. 29, No. 5, pp. 832–841, 2012.

- [7] Mike Allenspach, Yash Vyas, Matthias Rubio, Roland Siegwart, and Marco Tognon. Human-state-aware controller for a tethered aerial robot guiding a human by physical interaction. *IEEE Robotics and Automation Letters*, Vol. 7, No. 2, pp. 2827–2834, 2022.
- [8] LE Jane, LE Ilene, James A Landay, and Jessica R Cauchard. Drone & wo: Cultural influences on human-drone interaction techniques. In *CHI*, Vol. 17, pp. 6794–6799, 2017.
- [9] Jessica R Cauchard, Jane L E, Kevin Y Zhai, and James A Landay. Drone & me: an exploration into natural human-drone interaction. In *Proceedings of the 2015 ACM international joint conference on pervasive and ubiquitous computing*, pp. 361–365, 2015.
- [10] Parastoo Abtahi, David Y Zhao, Jane L E, and James A Landay. Drone near me: Exploring touch-based human-drone interaction. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, Vol. 1, No. 3, pp. 1–8, 2017.
- [11] Albert Mehrabian. Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in temperament. *Current Psychology*, Vol. 14, No. 4, pp. 261–292, 1996.
- [12] Rachel Kirby, Jodi Forlizzi, and Reid Simmons. Affective social robots. *Robotics and Autonomous Systems*, Vol. 58, No. 3, pp. 322–332, 2010.
- [13] Iolanda Leite, Carlos Martinho, and Ana Paiva. Social robots for long-term interaction: a survey. *International Journal of Social Robotics*, Vol. 5, No. 2, pp. 291–308, 2013.
- [14] Friederike Eyssel, Frank Hegel, Gernot Horstmann, and Claudia Wagner. Anthropomorphic inferences from emotional nonverbal cues: A case study. In *19th International Symposium in Robot and Human Interactive Communication*, pp. 646–651, 2010.
- [15] Paul Ekman and Wallace V Friesen. Constants across cultures in the face and emotion. *Journal of personality and social psychology*, Vol. 17, No. 2, p. 124, 1971.
- [16] Filippo Cavallo, Francesco Semeraro, Laura Fiorini, Gergely Magyar, Peter Sinčák, and Paolo Dario. Emotion modelling for social robotics applications: a review. *Journal of Bionic Engineering*, Vol. 15, No. 2, pp. 185–203, 2018.
- [17] Jekaterina Novikova and Leon Watts. A design model of emotional body expressions in non-humanoid robots. In *Proceedings of the second international conference on Human-agent interaction*, pp. 353–360, 2014.
- [18] James A Russell. A circumplex model of affect. *Journal of personality and social psychology*, Vol. 39, No. 6, p. 1161, 1980.
- [19] Cynthia Breazeal. Emotion and sociable humanoid robots. *International journal of human-computer studies*, Vol. 59, No. 1-2, pp. 119–155, 2003.
- [20] Meng-Ju Han, Chia-How Lin, and Kai-Tai Song. Robotic emotional expression generation based on mood transition and personality model. *IEEE Transactions on Cybernetics*, Vol. 43, No. 4, pp. 1290–1303, 2013.
- [21] Masahiro Mori, Karl F. MacDorman, and Norri Kageki. The uncanny valley [from the field]. *IEEE Robotics and Automation Magazine*, Vol. 19, No. 2, pp. 98–100, 2012.
- [22] Dante Tezza and Marvin Andujar. The state-of-the-art of human–drone interaction: A survey. *IEEE Access*, Vol. 7, pp. 167438–167454, 2019.
- [23] Dante Arroyo, Cesar Lucho, Silvia Julissa Roncal, and Francisco Cuellar. Daedalus: a suav for human-robot interaction. In *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction*, pp. 116–117, 2014.
- [24] Jessica R Cauchard, Kevin Y Zhai, Marco Spadafora, and James A Landay. Emotion encoding in human-drone interaction. In *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pp. 263–270. IEEE, 2016.
- [25] Chie Hieida, Hiroaki Matsuda, Shunsuke Kudoh, and Takashi Suehiro. Action elements of emotional body expressions for flying robots. In *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pp. 439–440. IEEE, 2016.
- [26] Peter AM Ruijten and Raymond H Cuijpers. If drones could see: Investigating evaluations of a drone with eyes. In *International Conference on Social Robotics*, pp. 65–74. Springer, 2018.
- [27] Robin Bretin, Mohamed Khamis, Emily Cross, and Mohammad Obaid. The role of drone’s digital facial emotions and gaze in shaping individuals’ social proxemics and interpretation. *ACM Transactions on Human-Robot Interaction*, Vol. 14, No. 3, pp. 1–34, 2025.
- [28] Yiyuan Wang, Luke Hespanhol, and Martin Tomitsch. How can autonomous vehicles convey emotions to pedestrians? a review of emotionally expressive non-humanoid robots. *Multimodal Technologies and Interaction*, Vol. 5, No. 12, p. 84, 2021.
- [29] Maria Franca, Nadia Bolognini, and Marc Brysbaert. Seeing emotions in the eyes: A validated test to study individual differences in the perception of basic emotions. *Cognitive Research: Principles and Implications*, Vol. 8, No. 1, p. 67, 2023.
- [30] Zheng Lian, Ya Li, Jian-Hua Tao, Jian Huang, and Ming-Yue Niu. Expression analysis based on face regions in real-world conditions. *International Journal of Automation and Computing*, Vol. 17, No. 1, pp. 96–107, 2020.
- [31] Sharath Chandra Guntuku, Mingyang Li, Louis Tay, and Lyle H Ungar. Studying cultural differences in emoji usage across the east and the west. In *Proceedings of the international AAIL conference on web and social media*, Vol. 13, pp. 226–235, 2019.
- [32] Gaku Kutsuzawa, Hiroyuki Umemura, Koichiro Eto, and Yoshiyuki Kobayashi. Classification of 74 facial emoji’s emotional states on the valence-arousal axes. *Scientific Reports*, Vol. 12, No. 1, p. 398, 2022.
- [33] Brigitte Fischer and Cornelia Herbert. Emoji as affective symbols: affective judgments of emoji, emoticons, and human faces varying in emotional content. *Frontiers in psychology*, Vol. 12, p. 645173, 2021.
- [34] Randy Gomez, Deborah Szapiro, Luis Merino, Heike Brock, Keisuke Nakamura, and Selma Sabanovic. Emoji to robomoji: exploring affective telepresence through haru. In *International Conference on Social Robotics*, pp. 652–663. Springer, 2020.
- [35] Unicode Consortium. Common Locale Data Repository (CLDR). <https://cldr.unicode.org/>, 2023. Accessed: 2025-08-02.
- [36] Saif Mohammad. Obtaining reliable human ratings of valence, arousal, and dominance for 20,000 english words. In *Proceedings of the 56th annual meeting of the association for computational linguistics (volume 1: Long papers)*, pp. 174–184, 2018.