Proceedings of the 24th IEEE International
Symposium on Robot and Human Interactive Communication
Kobe, Japan, Aug 31 - Sept 4, 2015

Design and Implementation of Multi-dimensional Flexible Antena-like Hair motivated by 'Aho-Hair' in Japanese Anime Cartoons: Internal State Expressions beyond Design Limitations

Kazuhiro Sasabuchi, Yohei Kakiuchi, Kei Okada and Masayuki Inaba

Abstract— **Recent research in psychology argue the importance of "context" in emotion perception. According to these recent studies, facial expressions do not possess discrete emotional meanings; rather the meaning depends on the social situation of how and when the expressions are used. These research results imply that the emotion expressivity depends on the appropriate combination of context and expression, and not the distinctiveness of the expressions themselves. Therefore, it is inferable that relying on facial expressions may not be essential. Instead, when appropriate pairs of context and expression are applied, emotional internal states perhaps emerge. This paper first discusses how facial expressions of robots limit their head design, and can be hardware costly. Then, the paper proposes a way of expressing context-based emotions as an alternative to facial expressions. The paper introduces the mechanical structure for applying a specific non-facial contextual expression. The expression was originated from Japanese animation, and the mechanism was applied to a real desktop size humanoid robot. Finally, an experiment on whether the contextual expression is capable of linking humanoid motions and its emotional internal states was conducted under a sound-context condition. Although the results are limited in cultural aspects, this paper presents the possibilities of future robotic interface for emotion-expressive and interactive humanoid robots.**

I. INTRODUCTION

Emotions for robots intend to improve the interaction capability between human and robot[1][2]. In order for humans to perceive robot emotions, recent researches focused on humanoid facial expressions and its ability to transmit emotional information. However, recent research in psychology argue that context (i.e., social situation, postures, voices) in emotion perception is essential, and that the facial expressions themself do not possess discrete emotional meanings[3]. Thus, facial expressions might not be the major modality for human-robot emotional communication, but rather one approach to represent a robot's internal state. In addition, facial designs of robots are restricted when we want to apply facial expressions. A mechanically facially expressive robot is either symbolic expressive (i.e., each facial part is seperated and directly actuated)[4][5] or realistic expressive (i.e., skin is actuated by moving mechanical parts underneath)[6][7]. A symbolic expressive face restricts robotic design to containing leaped out lips and mechanical shutter eyelids. A realistic expressive face may frighten and produce negative user reactions. Therefore, the design restricting and hardware costly facial expressions are, perhaps, not the optimum solution for expressing emotional internal states among humanoid robots. Instead, expressions that emphasize emotional meanings through certain context are perhaps suitable, especially for humanoids with small body size.

This paper first discusses how facial expressions limit robot hardware design, and the possibilities of non-facial contextual expressions in animation. The work is expected to contribute to ideas for humanoid robot head designs, and allow new methods for expressing their emotional internal states. Section II will be an overview of different types and approaches on robots' emotional expressions, including fictional ones. Section III discusses humanoid robots' facial expressions and the possible causes of its limitations. Section IV explains our robotic platform to test non-facial contextual expressions in animation. Section V to VII presents an experiment that tests the capabilities of non-facial contextual expressions. Section VIII is the conclusion.

II. BACKGROUND

There are several types of emotion or internal state related expressions that are implemented by actual robots. Body gestures such as neck and arm movements are possible to express emotions on robots with simple structure [8]. Generating human-like responses and gaze aversions allow a robot to be more thoughtful in a conversation [9]. Animal gestures, such as a dog's tail wag, also communicate emotional states when applied on a cleaning robot[10].

Other than expressions used by real-world humans or animals, there are also expressions such as ones applied to fictional characters. For example, Fig.1 represents an antenna-like hair movement("*aho-hair*") observed in Japanese animations such as "*monogatari series*" and "*saki*." The left image in the figure shows how a hair drawn in a question mark ("?") shape emphasizes a character's awareness. The right image in the figure represents how a deflated hair emphasizes a character's depressed emotion. Other expressions such as animating a specular reflective spot on a character's eye in a circular path, represents the character's feeling of expectation. Expressions such as robots with flashing lights are another example, as these could also be illustrated in fictional cartoons. The difference between robotic expressions and animation related expressions introduced above may be that robot-like expressions are distinct from human familiarity and may lack human-likeness. Animation related expressions

Department of Mechano-Informatics, University of Tokyo 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan. sasabuchi@jsk.imi.i.u-tokyo.ac.jp

are human-based, or at least human-based in the fictional world (that is, while the expressions are impossible by real world humans, they are possible by humans in the fictional world), and although such expressions are not intently possible as actual human movements, it focuses on human body parts. Therefore, robots with human-like appearance may be able to express their emotional state more effectively using animation expressions.

Fig. 1. Image of an antenna-like hair ("aho-hair") movement that emphasizes the feeling of a character in context. The expression is a typical expression in Japanese anime cartoons.

Recent researches in human robot interaction (HRI) have shown that the idea of using fictional expressions on robots yields remarkable results. Applying animation features to robots improves robot movements and allows human partners more accurately understand the robot's intent. Laws of Disney[11] create understandable robot movements in manipulation context. The anticipation rule, where a forethought action is conducted, improves readability of the robot's motion[12]. The exaggeration motion increases interaction detail retention[13]. These animation concepts over express human movements but nevertheless deduct human-likeness. Indeed, compared to the Disney laws, the contextual background of Japanese animation is strongly related to the Japanese culture and animation knowledge. This is true, although it must not be forgotten that the animation market itself holds a large population and has the potential of possessing familiarity to a large number of people.

Apart from the perception effects of animation based motions, [14] focused on the emotional effects of animation concepts. However, the emotional effects were evaluated by matching facial expressions with discrete emotions. This paper focuses on whether the use of animation concepts is effective in expressing emotions in a contextual manner. Thus, we evaluate emotional effects taking into consideration context, which is an important aspect as suggested by recent psychology literature.

III. ROBOT FACIAL EXPRESSIONS

This section gives a brief overview on mechanical facial expressions and how and why the design limitations and complicated high-DOF mechanical designs might occur. Although there are cost efficient and economical techniques for creating rich facial expressions, for instance 2D screen displays[15] or face projection[16], this paper will mainly focus on comparing mechanical approaches. There is a wide variety in the facial designs of existing robots, and the objective is to present new insights to robot facial design from a mechanical viewpoint. As explained in the introduction, robot facial expressions can be categorized as symbolic expressive and cartoon-like, or realistic expressive and real-human-like. While the appearances of robot belonging to these two categories differ, the fundamental mechanism is identical; manipulation of control points. Control points are points that shape the appearance of facial expressions. For example, a simple lip that either smiles or frowns necessitates three control points, one point to control the center of the lip, and one on each edge of the lip; stabilize the center point, then lower the two edge points to achieve a frown, or raise the two edge points to achieve a smile.

The manipulation of control points is similar to the manipulation of "vertices" in computer graphics(CG). Moving vertices creates facial expressions in computer graphics (this is called "morph targeting"). Though, the difference would be that while computer graphics "vertices" move freely in three-dimensional space, control points on robots are limited by hardware constraints. A control point on a realistic human-like robot such as [6], might possess a tendon-driven structure, thus the control point movement is limited in the driven direction. A control point on a cartoon human-like robot such as [5], might rotate on a single axis and movements are also limited in the rotational direction. Although with the use of additional actuator, a higher number of DOF can be achieved, most existing robots have at most 2 DOF per control point.

However, even with mechanical constraints, pseudo-three-dimensional facial movements are possible with 2-DOF control points. Fig.2 shows such an example of a mouth prototype. In the figure, the prototype opens its mouth widely in a way such that it is even noticeable from the side view. This is similar to expressions in animated cartoons or computer graphics cartoon characters. Despite the fact that vertices in computer graphics translate in three-dimensional space, the vertices are not randomly placed but are instead organized in a way that creates shapes on certain surfaces. For example, vertices that construct a mouth in 3D-CG space are also vertices that construct a character's face, and thus, must remain on the shape of the facial surface. A 2-DOF control point on a robot, that is constructed with similar surface constraints, are capable of expressing similar three-dimensional-alike facial movements. In other words,

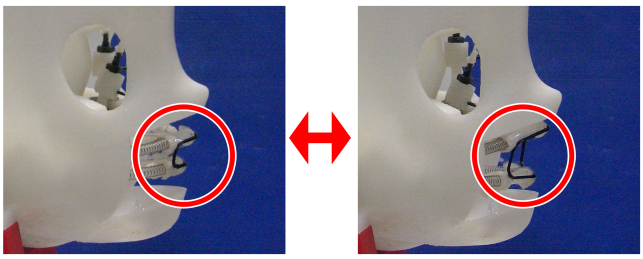

Fig. 2. Pseudo-three-dimensional facial movements using 3D-CG motivated surface constraints.

a robot's control point appears as if it has 3-DOF, if it mechanically slides across a 2-DOF rail, and if the rail is parallel to the robot's face-surface design. The face-surface design can vary, as long as the rails of control points remain parallel to the face surface.

Although the aforementioned concept allows more freedom in designing facial shapes, the hardware cost remains high. A simple mouth of this type contains a minimum of 10-DOF. Also, the face surface materials that would be controlled by the control points would be skins or magnet actuated exteriors, thus, the appearance of the robot is limited.

As explained above, facial expressions are costly. However, facial expressions have their own advantages of minimum cultural dependency. If handling complex mechanisms is not a problem, and the ability of showing various facial expressions is of high priority, then applying these hardware is certainly a valid choice for expressing emotions.

On the other hand, the proposed method in the next section is an alternative for expressing emotions for robots that are not capable of possessing such complex facial mechanism. The method proposed in the next section and the use of facial expressions may appear as conflicting choices for expressing emotions, but a combination of the two is another choice that may result in enhanced emotional expressivity.

IV. ROBOTIC PLATFORM

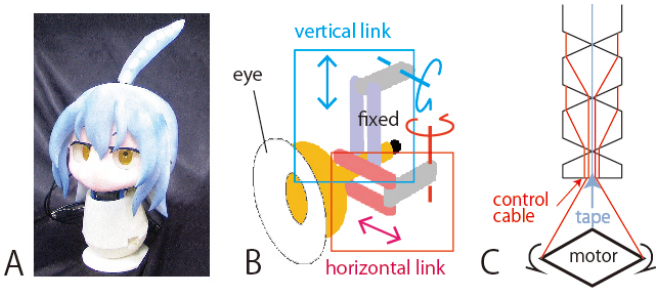

Fig. 3. Antenna-like hair implemented on an anime-like robot. A) Appearance of robot "sasabot". B) Concave eye mechanism. C) Mechanical structure of multi-dimensional flexible antenna-like hair motivated from Japanese anime expressions.

Our robot capable of animation context-based expressions is shown in Fig.3. The robot "sasabot" is designed with actuatable large eyes with concaved shape (Fig.3B). The advantage of concaved eyes is that it eliminates the spatial gap surrounding the eyes that a convex structure would otherwise create. According to feedback received at an exhibition, spatial gaps around the eye create an unease impression of the robot; therefore the elimination of such spatial gaps is required. The concave structure also allows us to create various eye shapes for the robot with the use of flat eyelids rather than convex eyelids. As shown in Fig.3A, "sasabot" has a half-moon shaped eye and colored customizable hair, which give personality to the robot. The robot achieves an animation-like appearance but does not employ any facial expressions. Instead, this robot operates an exterior antenna-like hair constructed with bilateral tendon-driven structure

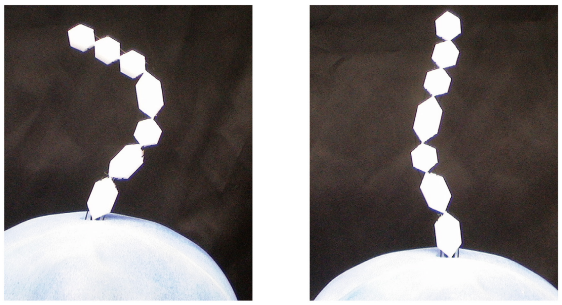

Fig. 4. Example of hair shapes expressed by the proposed mechanism.

and fiber tape-centered joints. Thus, the antenna-like hair maintains a miniature slender size. The hair is 12mm in diameter, 100mm at maximum length and has 7-DOFs. The fiber tape at the center is embedded during the 3D printing process of the whole mechanism, and connects the bead-like antenna joints. The control wire passing through the center of the antenna is rigid at full length but is flexible enough to bend and create expressional shapes. Two shapes that can be made by the mechanism are shown in Fig.4. Wrapping a silicon cover around completes the antenna-like hair.

This type of mechanism is likely to be implemented on humanoids that possess animation-like faces and animation-like hair. For general acceptance of the application, ear units for humanoids with antenna-like emotional feature are possible add-ons, such as shown in Fig.5. Humanoids with mechanical ear design are illustrated in cultural works and do not create a weird feeling (e.g. in "*Mahou Sensei Negima*" and "*Chobits*.") Baseline scores of the design of "sasabot" is shown in Fig.6. The data was collected at a science museum workshop from 9 elementary school children who played with the robot. The children rated each factor of cuteness, coolness, scariness, interest, human-likeness, and machine-likeness, as disagree, neither, or agree. In the figure, disagree is rated as 1, neither is rated as 2, and agree is rated as 3. Although the robot has a one-to-one head-to-body ratio, is desktop sized in height, and has an antenna-like hair, the robot received a 2.56 point rating for human-likeness and 1.11 point rating for scariness.

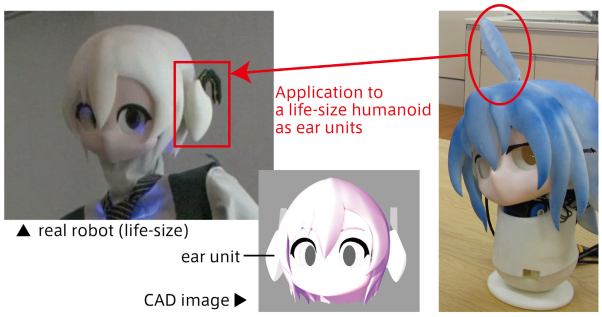

Fig. 5. Ear unit example of applying antenna-like emotional features on life-size humanoids. A possible and more general alternative instead of implementing hair.

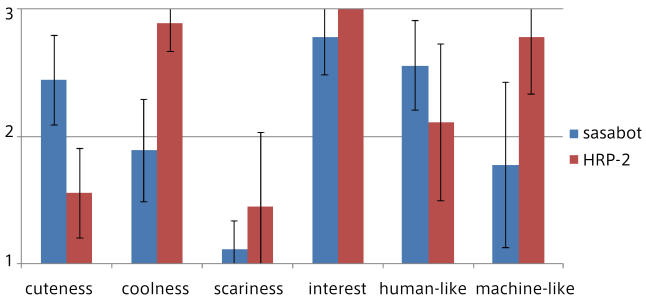

Fig. 6. Results of sasabot's baseline scores evaluated at a workshop. Error bars represent ±2SE. Scores 1 for disagree, 2 for neither agree or disagree, 3 for agree. Blue bars represent rated scores of sasabot, aside with comparative scores in red bars from the robot HRP-2, collected by the same participant group.

V. METHOD

A video-based laboratory experiment was conducted, where participants matched multiple sounds eith robot movements

The purpose of the experiment was to test whether a robot's non-facial expression was capable of communicating emotional internal states. As inferred from recent psychology studies, it is expected that when the expressions are under a contextual manner, they are capable of embodying emotional states. In other words, when a certain context is applied, the participants should perceive a robot's emotion through the robot's expressions. For non-facial expressions, operating the antenna hair-like mechanism introduced in the last chapter generated the expressional motions. For context, "sound" was selected. Sound is known to contain certain contexts. Results such as audiovisual bounce-inducing effect[17] imply that sound produces context to visual movements. In particular, the experiment focused on voice-recorded Japanese onomatopoeic words. Therefore, while onomatopoeia is not actually a sound but a word, we referred onomatopoeic words to create sound context.

Some related sound and emotion experiments evaluate emotion perception with musical sound, facial expressions, and a mixture of music and faces[18]. However, these experiments involve direct matching of expressions and particular emotional words (e.g. happy, sad, or angry), which is not practical considering the results of recent psychology. Matching posed and isolated expressions without context is unrealistic. To evaluate, observe, and analyze participant data, our experiment was also designed to have an identical matching scenario. However, to overcome the realistic issues, two alternative measures were implemented in our experiment. Firstly, instead of directly matching single motions (or already-paired contextual motions) with emotional words, the experiment was conducted following a two-step process. In the first step, participants matched contextual sounds with robot movements. In the second step, participants answered first-hand impressions of the contextual sounds. As a result, the motions and emotional meanings are not directly matched, but a contextually matched motion and its correlation to internal emotional messages are found. In this circumstance, motions are associated with emotional meanings if and only if correct context is applied, which is, according to psychological theories, the case in reality.

The second measure for overcoming realistic issues, relates to the vagueness of emotional internal states. While some people might percieve an expression as grief others might percieve the identical expression as disgust. In order to prevent biasing the participants' answers by asking leading questions with discrete emotional words, a numerical scale indicating positive-negative or strong-weak impression was evaluated for each sound. Therefore, the emotional internal state evaluated in the experiment represented a strongly positive internal state, a strongly negative internal state, a weakly positive internal state, a weakly negative internal state, or a neutral state. If the participants were able to match a contextual sound and movement, and if the contextual sound scored a strongly positive state, the movement when attached to the sound context, is regarded as an action obtaining strongly positive emotions.

The experiment was conducted with 10 Japanese participants of ages 20 to 29. Each participant observed a total of three videotaped robot movements of the antenna-like animation expressive hair. In order to control the experiment and have participants concentrate on the robot's actions, we removed the environmental noises and motor sounds from the videos. For each movement, the participants were asked to choose the most appropriate sound out of five spoken onomatopoeic words or answered none if none of the choices were thought to be suitable. The participants were also asked to answer any possible choices other than the most matched onomatopoeia. The possible choice question would explain the results of the most matched onomatopoeia. If the most matched selection varies, the possible choice question will distinguish whether the results were due to existence of onomatopoeia of similar emotional context, or whether the movement did not possess a clear emotional context. After matching the three movements and sounds, the participants rated the positive-negative and strong-weak impression of each sound. Fig.7 shows how the participant matched the spoken onomatopoeic words with robot motions using a movie editing software.

Fig. 7. Picture of how the experiment was conducted. Participants selected the most matched voice for each motion by operating a movie editor software on a computer screen.

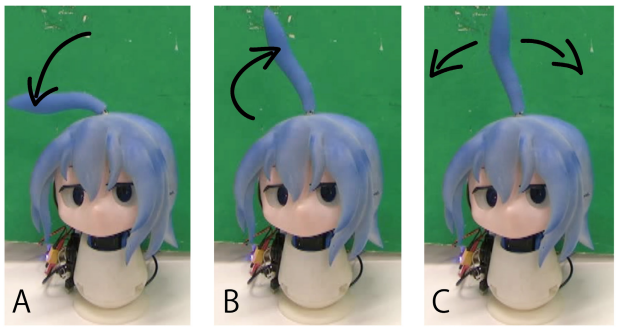

Fig. 8. The three movements tested. A) Waving down movement. B) Sudden wake-up movement. C) Side by side movement.

The participants chose which onomatopoeic word to apply by turning on and off each sound aligned in the software, and then clicking the play button to check if the sound matched the motion. The play duration of the onomatopoeic words was organized so that all the onomatopoeic words played for about the same time duration.

The three evaluated movements are illustrated in Fig.8. In A) the antenna-like hair waved down with a slow-in timing. In B) the antenna-like hair presented a sudden wake up movement like an energetic spring. In C) the antenna-like hair waved side by side quickly.

The contextual sound (spoken onomatopoeic words) selected in the experiment was "*ha*," "*iraira*," "*gagan*," "*mumu*," and "*shobon*." The five sounds are all Japanese onomatopoeia. "*ha*" is a sound that expresses a startled context. "*iraira*" is a sound that expresses a frustration context. "*gagan*" is a sound that expresses a incredibly shocking or depressed context. "*mumu*" is a sound that expresses at a loss for words context. "*shobon*" is a sound that expresses a discouraged context.

VI. RESULTS

Results of the selected most matched sound for each motion are shown in Fig.9A. 80% of the participants for the waving down, 60% for the sudden wake up, 50% for the waving side-by-side motion selected an identical onomatopoeic word. Therefore, more than half of the participants linked a typical contextual sound for each motion.

Selected sounds including the possible choice responese are shown in Fig.9B. The results indicate that possible context may vary, but more than 40% chose the same answer.

The positive-negative impression of each sound is shown in Fig.9C and the strong-weak impression in Fig.9D. 0.0 is the minimum score and 1.0 is the maximum. 0.0 is negative and weak, 1.0 is positive and strong. Participants marked a score on a solid line and the marked position was converted to a 0.0 to 1.0 scale. Taking into account these impression data and the possible context data, 73% of the selections for the waving down movement scored an average of less than 0.29 in strong-weak impression.

Thus, while the context selection varied between "*mumu*" and "*shobon*," the emotional internal state expressed by the waving down movement was a weak expression. For the

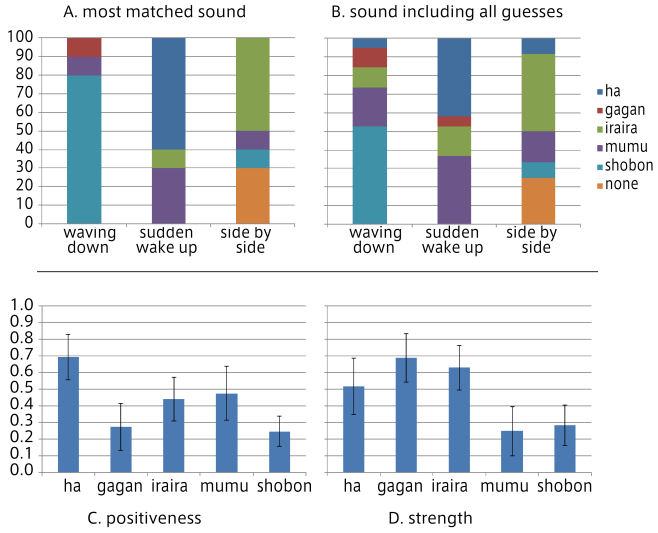

Fig. 9. Experiment results. A) Selected most matched sound for the three robot motions. B) Selected sound including all possible guesses for the three robot motions. C) Positive/negative impression score of each sound context. D) Strong/weak impression score of each sound context. Error bars ±2SE.

sudden wake up movement, only 5.3% of the selections scored an average under 0.30 in positive-negative impression. The emotional internal state of sudden wake up thus does not follow a negative impression, and considering the most matched contextual sound, may include a positive internal emotion. For the side by side movement, 30% selected that none of the context matched the movement, and a tendency for internal emotion were not observed from the selected all possible choices.

VII. DISCUSSION

The results support the idea of non-facial contextual expressions representing an emotional internal state for humanoid robots. The specific expression measured in the study was motivated by Japanese animation, and a multi-dimensional flexible hair behaved in three ways. The results present that waving down and sudden wake up motions matched contextual sounds, and was capable of representing emotional internal states of weak and positive.

This work was limited to Japanese onomatopoeic words, hair movements observed in Japanese animation, and Japanese participants. Therefore, conclusions on the perception of contextual emotional meanings in other countries cannot be drawn from the results of this experiment. Cultural aspects and cultural context should affect a human's perception of a robot's motion.

However, overcoming cultural differences may not always be a necessary. It may be an issue for service robots that are serving people from different countries, but considering that even manner styles differ among people of the same culture, identifying a robot's motion is not a substantial issue for personal robots. Instead, motions that contain acceptance, reliability, and compatibility at an individual or group level are more of a concern. During the experiment, some participants easily selected a matching sound and showed definite confidence, while other participants answered that

neither sound matches for the same motion. Other noticeable behaviors were that participants had definite confidence even while selecting a different sound compared to the selection of majority. Difference in knowledge and backgrounds of participants were apparent, yet each individual found preferable motions in corresponding contexts. A calibration system of motion preference may improve acceptance and interaction with a robot. To achieve the reliability and compatibility goals, emphasizing internal states and expressing motions in correct contextual manners is a valid method. Our study demonstrates that while some sound contexts match certain motions, others do not. A non-matching motion and context is possibly misleading and can confuse a human's understanding of a robot's internal state. In contrast, a well-employed motion in context can emphasize its internal meanings.

Further study on identifying the type of contextual expressions that is effective among cultures other than the Japanese culture, or groups unfamiliar with Japanese animation or related entertainment fields, is to be studied. In addition, how the hair movement is applicable to other robot parts, such as the proposed ear unit, is an important direction for future work.

VIII. CONCLUSION

Context in emotion perception is an important factor according to recent studies in psychology. While, facial expressions do not possess discrete emotional meanings, expressions under correct context shows internal states.

This paper shows how applying a non-facial expression of multi-dimensional flexible antenna-like hair movement, is internal state expressive with onomatopoeia sound context. The expression was motivated by Japanese anime cartoons, and was tested on a desktop size real humanoid "sasabot."

The proposed contextual expression had the capability of linking motions and emotional internal states, such as a waving down movement expressing weak emotions and a sudden wake up movement expressing positive emotions. Although limited to Japanese culture, the expression's hardware is of low cost and the structure size is implementable on small robots. This overcomes the design limitations that implementing facial expressions will otherwise encounter. The robot's design with the proposed expression was marked as cute, and although not a expression directed by actual human and only fictional characters, the robot's appearance was scored as human-like.

Humanoid robots are not real humans. They may possess human-like appearance, they may behave similar to humans for interaction understandability, but because they are machines, they should not be limited to having identical expressions as human. Exaggerated expressions that surpass human capabilities are applied to human-like characters in the fictional world of animation. In order to achieve better emotional communication with robots, it may be important to consider different modality beyond human capabilities, such as the ones observed in Japanese animation. Nevertheless, the key to selecting an expression is to understand the possibility of its expressivity under a contextual manner. Emphasizing and directing the perception of a robot's internal state should not be separated from the context.

REFERENCES

[1] Cynthia Breazeal and Rodney Brooks. Robot emotion: A functional perspective. *Who needs emotions*, pp. 271–310, 2005.

[2] Hyoung-Rock Kim, KangWoo Lee, and Dong-Soo Kwon. Emotional interaction model for a service robot. In *Robot and Human Interactive Communication, 2005. ROMAN 2005. IEEE International Workshop on*, pp. 672–678. IEEE, 2005.

[3] Lisa Feldman Barrett, Batja Mesquita, and Maria Gendron. Context in emotion perception. *Current Directions in Psychological Science*, Vol. 20, No. 5, pp. 286–290, 2011.

[4] I Lutkebohle, Frank Hegel, Simon Schulz, Matthias Hackel, Britta Wrede, Sven Wachsmuth, and Gerhard Sagerer. The bielefeld anthropomorphic robot head "flobi". In *Robotics and Automation (ICRA), 2010 IEEE International Conference on*, pp. 3384–3391. IEEE, 2010.

[5] Gabriele Trovato, Tatsuhiro Kishi, Nobutsuna Endo, Kenji Hashimoto, and Atsuo Takanishi. Development of facial expressions generator for emotion expressive humanoid robot. In *Humanoid Robots (Humanoids), 2012 12th IEEE-RAS International Conference on*, pp. 303–308. IEEE, 2012.

[6] Takuya Hashimoto, Sachio Hitramatsu, Toshiaki Tsuji, and Hiroshi Kobayashi. Development of the face robot saya for rich facial expressions. In *SICE-ICASE, 2006. International Joint Conference*, pp. 5423–5428. IEEE, 2006.

[7] Ho Seok Ahn, Dong-Wook Lee, Dongwoon Choi, Duk-Yeon Lee, Manhong Hur, and Hogil Lee. Appropriate emotions for facial expressions of 33-dofs android head ever-4 h33. In *RO-MAN, 2012 IEEE*, pp. 1115–1120. IEEE, 2012.

[8] Jamy Li and Mark Chignell. Communication of emotion in social robots through simple head and arm movements. *International Journal of Social Robotics*, Vol. 3, No. 2, pp. 125–142, 2011.

[9] Sean Andrist, Xiang Zhi Tan, Michael Gleicher, and Bilge Mutlu. Conversational gaze aversion for humanlike robots. In *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction*, pp. 25–32. ACM, 2014.

[10] Ashish Singh and James E Young. Animal inspired peripheral interaction. *Peripheral Interaction: Embedding HCI in Everyday Life*, p. 33.

[11] Frank Thomas, Ollie Johnston, and Frank. Thomas. *The illusion of life: Disney animation*. Hyperion New York, 1995.

[12] Leila Takayama, Doug Dooley, and Wendy Ju. Expressing thought: improving robot readability with animation principles. In *Proceedings of the 6th international conference on Human-robot interaction*, pp. 69–76. ACM, 2011.

[13] Michael J Gielniak and Andrea L Thomaz. Enhancing interaction through exaggerated motion synthesis. In *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction*, pp. 375–382. ACM, 2012.

[14] Tiago Ribeiro and Ana Paiva. The illusion of robotic life: principles and practices of animation for robots. In *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction*, pp. 383–390. ACM, 2012.

[15] Hirotaka Osawa. Emotional cyborg: complementing emotional labor with human-agent interaction technology. In *Proceedings of the second international conference on Human-agent interaction*, pp. 51–57. ACM, 2014.

[16] Frédéric Delaunay, Joachim De Greeff, and Tony Belpaeme. Towards retro-projected robot faces: an alternative to mechatronic and android faces. In *Robot and Human Interactive Communication, 2009. RO-MAN 2009. The 18th IEEE International Symposium on*, pp. 306–311. IEEE, 2009.

[17] Massimo Grassi and Clara Casco. Audiovisual bounce-inducing effect: When sound congruence affects grouping in vision. *Attention, Perception, & Psychophysics*, Vol. 72, No. 2, pp. 378–386, 2010.

[18] Eun-Sook Jee, Chong Hui Kim, Soon-Young Park, and Kyung-Won Lee. Composition of musical sound expressing an emotion of robot based on musical factors. In *Robot and Human interactive Communication, 2007. RO-MAN 2007. The 16th IEEE International Symposium on*, pp. 637–641. IEEE, 2007.