Emotional Expression Humanoid Robot WE-4RII
-Evaluation of the perception of facial emotional expressions by using fMRI-

M. Zecca[1,2], T. Chaminade[3], M.A. Umiltà[4], K. Itoh[2,5], M. Saito[6], N. Endo[6], Y. Mizoguchi[6], S. Blakemore[3], C. Frith[3], V. Gallese[4], G. Rizzolatti[4], S. Micera[7], P. Dario[2,7], H. Takanobu[2,8,9], A. Takanishi[1,2,5,9,10]

1. Consolidated Research Institute for Advanced Science and Medical Care, Waseda University, Tokyo, Japan, email: zecca@aoni.waseda.jp; takanisi@waseda.jp
2. RoboCasa, Waseda University, Tokyo, Japan
3. Wellcome department of imaging neuroscience University College of London
4. Dipartimento di Neuroscienze, Sezione di Fisiologia, Università di Parma
5. Department of Mechanical Engineering, Waseda University, Tokyo, Japan
6. Graduate School of Science and Engineering, Waseda University, Tokyo, Japan
7. ARTS Lab, Scuola Superiore Sant'Anna, Pisa, Italy
8. Department of Mechanical Systems Engineering, Kogakuin University, Tokyo, Japan
9. Humanoid Robotics Institute (HRI), Waseda University, Tokyo, Japan
10. Advanced Research Institute for Science and Engineering, Waseda University, Tokyo, Japan

Personal robots and robot technology (RT)-based assistive devices are expected to play a major role in our elderly-dominated society, with an active participation to joint works and community life with humans. In order to achieve this smooth and natural integration between humans and robots, interaction also at emotional level is a fundamental required.

Objective of this research, therefore, is to clarify how the emotions expressed by a humanoid robot are perceived by humans. The preliminary results show several similarities but also several differences in perception.

***Key Word*s**: facial emotional expression, humanoid robot, fMRI

1 Introduction

Japan has the world's highest percentage of senior citizens over 65 (21%) and the smallest percentage of children under 15 (13.6%) [1]. These figures show that Japanese society is aging much faster than expected, and they underscore the effects of a shrinking birthrate [2]. In this aging society, it is expected that there will be a growing need for home, medical and nursing care services, including those provided by robots, to assist the elderly both on the physical and the psychological levels [3]. In this regard, human-robot communication and interaction are very important, particularly in the case of home and personal assistance for elderly and/or handicapped people. If a robot had a "mind" (intelligence, emotion, and will) similar to the human one, it would be much easier for the robot to achieve smooth and natural adaptation and interaction with its human partners and the environment [4].

Takanishi et al have been developing developed the WE-3 (Waseda Eye No.3) series since 1995. So far they have achieved coordinated head-eye motion with V.O.R. (Vestibular-Ocular Reflex), depth perception using the angle of convergence between the two eyes, adjustment to the brightness of an object with the eyelids and four senses, visual, auditory, cutaneous and olfactory sensations. In addition, they obtained the expression of emotions by using not only the face, but also the upper-half of the body with the Emotion Expression Humanoid Robot WE-4 (Waseda Eye No.4) series with the waist, 9-DOFs emotion expression humanoids arms and humanoid robot hands RCH-1 (Robo Casa Hand No.1) [5-7].

WE-4RII transmission of emotions was evaluated by showing the movies of its six basic emotional expressions exhibited to many subjects. The users chose the emotion they thought the robot expressed. The averaged recognition rate of all emotional expressions of WE-4RII was 93.5 [%], which proved that WE-4RII can effectively convey its emotions using its upper-half bodily expressions [7].

However, this kind of analysis lacks of objectivity. In order to obtain more objective data about the user perception of the emotions, a different approach should be pursued.

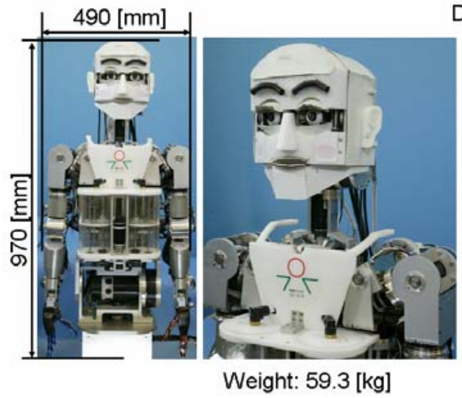
The mirror neuron system [8] is an area of our brain whose neurons fire both when we perform an action and when we observe the same action performed by someone else. The function of the mirror system is a subject of much speculation. These neurons may be important for understanding the actions of other people, and for learning new skills by imitation. It is also considered that the Mirror Neuron System plays an important role in the recognition of emotions.

Objective of this research, therefore, is to clarify how the emotions expressed by a humanoid robot are perceived by humans.

2 Material and Methods

2.1 Emotion Expression Humanoid Robot WE-4RII

The Emotion Expression Humanoid Robot WE-4RII (see Fig. 1) developed in Takanishi lab is capable of expressing 6 different emotions (Happyness, Anger, Surprise, Sadness, Disgust, Fear) by using facial expressions and movements of the neck, the arms and the hands [7].



Degrees of Freedom

Part	DOF
Neck	4
Eyes	3
Eyelids	6
Eyebrows	8
Lips	4
Jaw	1
Lung	1
Waist	2
Arms	18
Hand	12
Total	59

Fig. 1 The Emotion Expression Humanoid Robot WE-4RII.

In Fig 2 the details about the mechanisms used to obtain different facial expressions are presented. Eyebrows are realized by sponge. Each eyebrow is actuated by 4 DC motors connected by clear wire. Lips are obtained by 2 spinlde shaped springs. Their movement is realized by 4 DC motors. Eyelids have 6 DOFs.

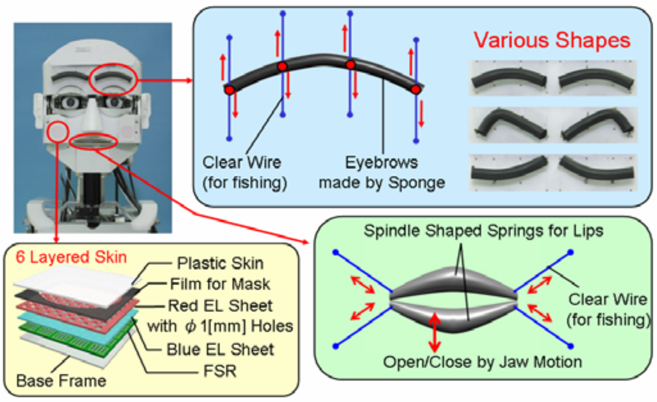


Fig. 2 Details of the mechanisms for facial expressions.

2.2 Experimental Paradigm

The experimental paradigm contained 16 conditions defined by a 2x2x4 factorial design with factors:

- 2 Agents: human or humanoid robot WE-4RII;
- 2 Identities of each agent: bald / hairy version;
- 4 facial motions depicted by the agent: silent speech (articulating a syllable e.g. "ba ba") or emotion (happiness, anger or disgust).

Silent speech was selected because it is supposed to activate the motor areas while not activating the emotional area of the Mirror Neuron System.

Each stimulus consisted of 1.5 seconds greyscale videoclips (i.e. 38 frames at 25 frames per second). One example for happyness for one human actor and for the robot with the wig is presented in Figure 3.

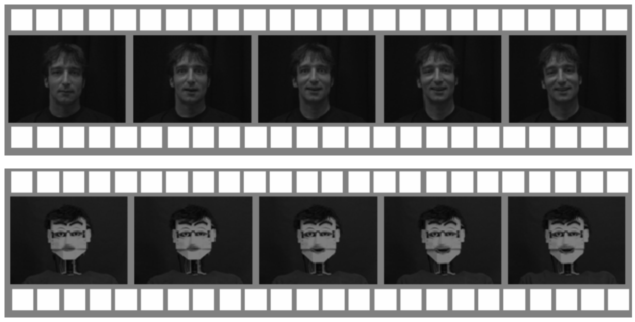


Fig.3 Example of Happyness for one human actor (top) and for the robot WE-4RII with the wig (bottom)

Two different actors were recorded to prepare the human stimuli while two versions of the robotic face were prepared by the addition of a wig (Figure 4).

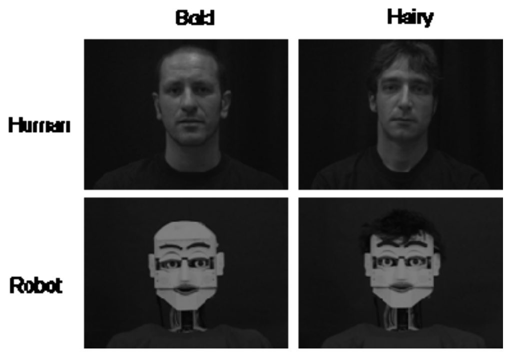


Fig. 4 2x2 Factorial Plan

Four different versions of each type of stimuli were used (see fig 5 for happiness). All stimuli started from a neutral pose and stopped with the emotional expression. Great care was taken to match the dynamics of the human and robot stimuli pairwise as much as possible, to minimize the false responses. The greyscale was digitally modified to match the background colour and the overall contrast between the human and robot stimuli. The overall luminosity of the clips was reduced to avoid too much visual fatigue to the subjects under fMRI.

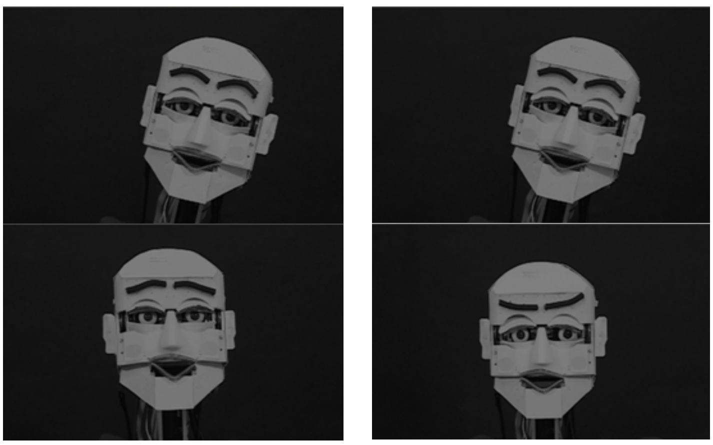


Fig. 5: four different samples for Happiness.

2.3 Behavioral Analysis

Participants were asked to recognize the emotional content of the stimuli. After presentation of emotional stimuli, subjects had to chose between 4 emotions in a forced choice paradigm (Angry, Happy, Disgust, Neutral). There were 8 blocks of 8 stimuli for each participant reported here, and each stimulus was shown once. All subjects experienced 1 to 4 experimental sessions before the one used for the present analysis, ensuring they all experienced the robot and the stimuli at the time of acquisition of the behavioral data reported here.

2.4 fMRI acquisition

The fMRI acquisition consisted of 4 sessions, each one composed by 8 blocks (4/emotional, 4/neutral tasks), with 8 stimuli per block (1 of each type).

Each presentation was followed by 1.5 seconds response screen, with Stimulus-onset asynchrony (SOA) jittered (normal distribution, 6±0.7 seconds), divided before and after stimulus.

Reminder of task, continuous Likert scale [-200 200] with target emotion and “None” for emotional task, “Lots” and “None” for neutral task, side random

Standard SPM2 analysis with specific EPI sequence & unwarp with fieldmaps was used.

fMRI of the whole brain, with a sequence optimized for amygdala, orbitofrontal and ventrotemporal cortex, composed by 48 slices, 3x3x3mm^3, TR=4.32 secs, @1.5T was used. Each subject worked for 1 hour.

Two different types of question were asked to the subjects:

1. Emotional: “How emotional was the face?” Rating from “Neutral” to the target emotion (i.e. “Happy”);
2. Neutral: “How much movement did the face show?” From “None” to “Lots”

3 Experimental Results

3.1 Behavioral Analysis

10 subjects participated after giving their informed consent, independently of or in addition to the fMRI experiment. One subject did not report any stimulus as "Neutral", and was removed from the analysis, so that n=9 for the results reported here.

Analysis of variance (factors of interest: Emotion, Agent; random factors: version of the agent [bald vs hairy], subjects) indicates a significant main effect of the agent used to display the emotion on the ratio of correct answers (number correct divided by total number for each condition), but no significant effect of the emotion on the recognition or interaction between the emotion and the agent.

Results of the behavioral analysis are presented in Table 1 and Figure 6. The difference between human and robot agents is highly significant (t=4.512, p<0.001), with emotions being better recognized for the human (98%) than for the robot (85%) agents.

Table 1. Recognition ratio depending for each agend and emotion.

Agent	Emotion	Mean	Std. Deviation
Human	Angry	1.0000	.00000
	Disgust	.9722	.11785
	Happy	.9722	.08085
	Neutral	.9722	.11785
	Total	.9792	.09145
Robot	Angry	.8333	.32084
	Disgust	.9167	.21004
	Happy	.7639	.23440
	Neutral	.8750	.21437
	Total	.8472	.25020

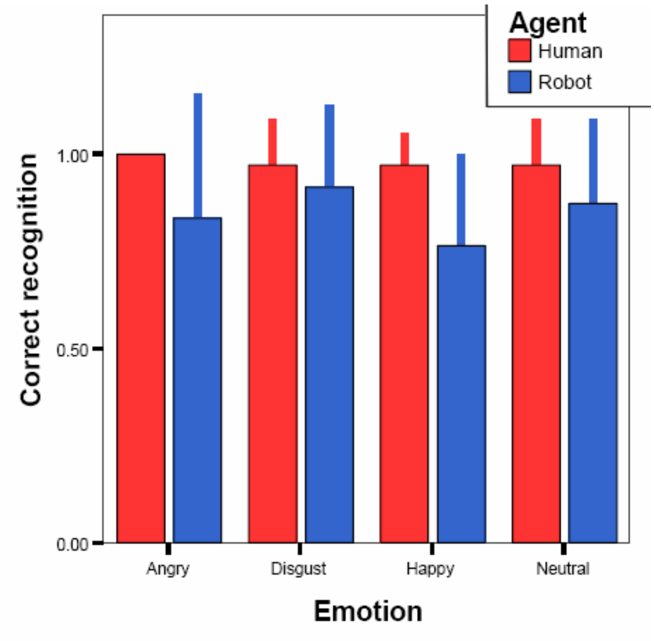


Fig 6. Recognition ratio depending on each agent and emotion.

3.2 fMRI analysis – behavioral analysys

13 subjects (Male: 4; Femail:9), all right handed, Average age: 29.4 ±7, range 22.4 – 39.7, gave their informed consent to the participation to this experiment.

One-way ANOVAs restricted to each emotion were used to assess differences in ratings due to the agent used to depict the emotion. Only emotional rating of the Angry stimulus was significantly different (p<0.001).

Human more emotional and perceived as moving more than robot across all conditions. It could be either subjective or due to stimuli not being perfectly matched, though.

Only emotional ratings of Angry stimuli are significantly affected by the agent (see fig. 7 and 8 for reference). Therefore, angry stimuli will be excluded from the fMRI analysis.

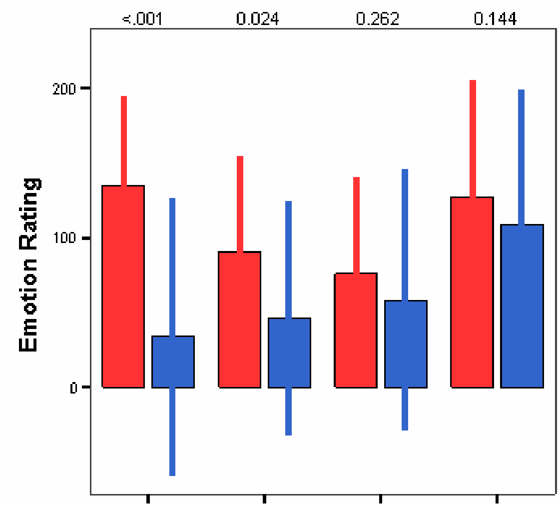


Fig 7. Emotion ratings (error bar: SD) depending to agent and emotion. Number on top gives the effect of agent according to ANOVA for each emotion and rating.

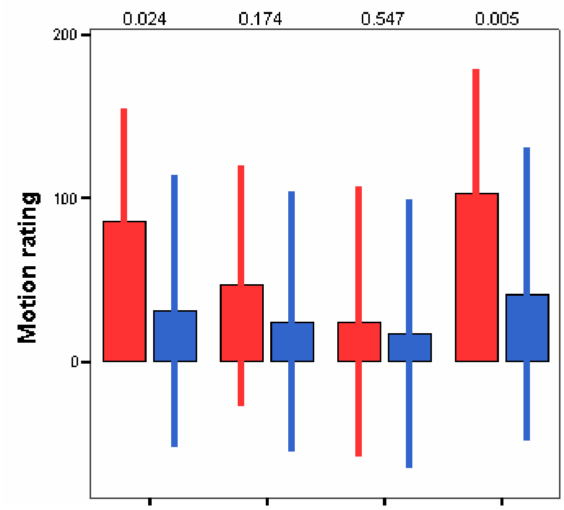


Fig. 8: Motion ratings (error bar: SD) depending to agent and emotion. Number on top gives the effect of agent according to ANOVA for each emotion and rating.

3.3 fMRI analysis – further analysys

At the time of the compilation of this paper, further analysis of the data is still in progress. The results will be published in a following paper.

4 Discussion and Conclusion

Objective of this work was to clarify how the emotions expressed by a humanoid robot are perceived by humans. To do this, we analyzed the response of the Mirror Neuron System of several subjects looking at videos of the robot and of human actors, and we compared the responses.

A pure behavioral analysis showed that the difference between human and robot agents is highly significant, with emotions being better recognized for the human (98%) than for the robot (85%) agents (see Fig 6 and Table 1).

The analysis of the fMRI data is still preliminary. However, some simple conclusion can be drawn. In a one-way ANOVAs restricted to each emotion, only emotional rating of the Angry stimulus was significantly different ($p<0.001$).

In the future, the analysis will be extended to different goups of subjects, in order to clarify the dipendencies with age, sex, and cultural backgound. The analysis will be also extended to full body emotional expressions, which is believed to be even more important than facial expressions [9] (i.e.. while a fearful faces signal a threat, it does not provide information about either the source of the threat or the best way to deal with it. By contrast, fearful body positions signal a threat and at the same time specify the action undertaken by the individuals fearing for their safety).

ACKNOWLEDGMENT

Part of this research was conducted at the Humanoid Robotics Institute (HRI), Waseda University. The authors would like to express thanks to Okino Industries LTD, OSADA ELECTRIC CO., LTD, SHARP CORPORATION, Sony Corporation, Tomy Company LTD and ZMP INC. for their financial support for HRI. And, the authors would like to thank Italian Ministry of Foreign Affairs, General Directorate for Cultural Promotion and Cooperation, for its support to the establishment of the ROBOCASA laboratory. In addition, this research was supported by a Grant-in-Aid for the WABOT-HOUSE Project by Gifu Prefecture. Part of the research has been supported by the EU FET NEUROBOTICS FP6-IST-001917 "The fusion of Neuroscience and Robotics". Finally, the authors would also like to express thanks to ARTS Lab, NTT Docomo, SolidWorks Corp., Consolidated Research Institute for Advanced Science and Medical Care, Waseda University, Advanced Research Institute for Science and Engineering, Waseda University, Prof. Yutaka Kimura, Dr. Yuichiro Nagano and Dr. Naoko Yoshida for their support for our research.

References

[1] Ministry of Internal Affairs and Communications of Japan, Preliminary Report on the 2005 National Census, June 2006.

[2] NIPS. Population statistics of Japan 2003. Technical report, National Institute of Population and Social Security Research, Japan, 2003

[3] JApan Robot Association, Summary Report on Technology Strategy for Creating a Robot Society in the 21st Century, May 2001.

[4] S. Hashimoto, "KANSEI as the Third Target of Information Processing and Related Topics in Japan." Proceedings of the International Workshop on KANSEI: The Technology of Emotion, pp.101-104., 1997.

[5] H. Miwa, K. Itoh, et al.: "Design and Control of 9-DOFs Emotion Expression Humanoid Arm", Proceedings of the 2004 IEEE International Conference on Robotics and Automation, pp.128-133, 2004.

[6] H. Miwa, K. Itoh, et al.: "Effective Emotional Expressions with Emotion Expression Humanoid Robot WE-4RII", Proceedings of the 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp.2203-2208, 2004.

[7] M. Zecca, A. Takanishi, et al, "On the development of the Emotion Expression Humanoid Robot WE-4RII with RCH-1", HUMANOIDS 2004, v1, Page(s):235 - 252.

[8] G. Rizzolatti, L. Craighero, "The Mirror-Neuron System",Annual Review of Neuroscience, Vol. 27: 169-192, July 2004

[9] B. De Gelder, "Towards the neurobiology of emotional body language", Nature, v7:2006, pp 242-249