How does the robot feel? Annotation of emotional expressions generated by a humanoid robot with affective quantifiers

Mina Marmpena[1], Angelica Lim[2], and Torbjørn S. Dahl[1]

***Abstract*— Human-robot interaction could be greatly enhanced if we understand and improve the reliability and impact of robot emotional expressions. Using a pre-designed set of robot animations as our starting point, we seek to increase its usability by annotating it with valence and arousal quantifiers. An initial experiment is described that aims to provide such an annotation by evaluating the quality and consistency of human emotional interpretation of the robot animations.**

I. Introduction

The integration of emotional traits in human-robot interaction (HRI) could enhance user experience and establish empathic communication and collaboration between humans and robots. Humanoid robots are increasingly populating social environments like homes, schools and commercial facilities. As a consequence, their ability to interact with humans in a natural way is becoming increasingly important [1]. Body language and general non-verbal cues like gestures and non-linguistic speech are indispensable features of empathic communication. Work has been conducted on key body postures and head positions [2], showing that such cues can be interpreted in meaningful ways and used to improve the expressiveness of robots. However, the effects of emotional expressions in the form of animations, have not yet been widely studied.

To evaluate such effects, this work deployed dimensional theories of emotion, more specifically the Circumplex Model [3] as a representative of non-directional core affect. According to this approach, emotional states can be represented on a two-dimensional surface defined by a *valence* (pleasure/displeasure) and an *arousal* axis. These two measures are widely accepted as quantitative descriptors of emotion and have the added advantage of describing a continuous mapping that is potentially more effective in capturing subtle differences than measures based on discrete emotion theories that aggregate many different expressions in a set number of categories.

The study presented uses a set of pre-designed animations for a humanoid robot, the Pepper robot from SoftBank Robotics which was created in a category-oriented way to convey emotions [4]. The study aims to map these animations onto a two-dimensional space of affect. First, it aims to study how humans perceive non-verbal emotional expressions executed by a humanoid robot, to establish whether the expressions are distinguishable by human observers, and to evaluate the consistency of their interpretations. Second, the study aims to generate quantitative affect descriptors (valence and arousal), that can provide a high-resolution tool that is more sensitive than discrete categorical tags. Such a tool could be potentially be re-used by other researchers and robotic application developers.

[1]Mina Marmpena and Torbjørn Dahl are with the Centre for Robotics and Neural Systems, School of Computing, Electronics and Mathematics, University of Plymouth, Plymouth, UK asimina.marmpena|torbjorn.dahl@plymouth.ac.uk

[2]Angelica Lim is with SoftBank Robotics Europe, Paris, France alim@aldebaran.com

II. Methods and Materials

A. Participants

The experiment proposed has been granted ethical approval by the University of Plymouth, Faculty of Science and Engineering Research Ethics Committee. The participants will include 20 adults aged between 18 and 35. The group will be balanced in terms of gender and subjects will have a minimal or no previous experience with the Pepper robot.

B. Hardware

The study will use the Pepper robot platform, a humanoid robot created by SoftBank Robotics. The Pepper robot is 120cm tall and weighs 28 kg. It has 20 degrees of freedom, including a head, two arms and a wheeled base. The operating system running on the robot is NAOqi 2.5.5.5 and the experiment uses the Python SDK to control robot's actions. The participant interface, used for data collection, is a 19.5-inch tablet. The interface is implemented as a web application with Django 1.11.

C. Animations

The set of the robot's emotional expressions consists of 36 animations designed by SoftBank Robotics animators to convey different categories of emotion. Each animation lasts for approximately 10 seconds and consists mainly of body motion (not locomotion), although some of the animations make use of additional interactive modalities including eye LEDs patterns, sounds or both. The selection of 36 animations was done from a broader set to reflect emotions in nine different discrete valence/arousal value combinations as presented in Table I.

TABLE I
VALENCE/AROUSAL VALUE COMBINATIONS

Negative/Excited	Neutral/Excited	Positive/Excited
Negative/Calm	Neutral/Calm	Positive/Calm
Negative/Tired	Neutral/Tired	Positive/Tired

D. Ratings, scales and questionnaires

The study uses the *Affective Slider* [5], a digital tool that has been validated as a reliable scale for the quick measurement of the emotion dimensions of valence and arousal. The main advantage of the Affective Slider is that we can collect the ratings on continuous scales (we use a resolution of 100 points). Furthermore, it is a rather intuitive interface and does not require significant training. The slider is integrated into a user interface along with a play button for the participants to control the robot.

The experiment will make use of an affect measuring questionnaire [6] ahead of the subject's interaction with the robot in order to indicate the subject's positive and negative mood at the time of the experiment. Furthermore, two five-point Likert scales will be presented for the participants to indicate their level of confidence in the arousal and valence ratings they submit for each trial.

E. Experimental Procedure

After registering the demographic data (age, sex, nationality), the participants will be seated in front of the robot at a distance of 2 meters. First they complete the affect measuring questionnaire, and then they go through a training session of 3 trials. Each trial requires the participant to observe and rate a single animation. The participant response is solicited by the question *"How does the robot feel?"*. When the participants indicate that they understand the distinction between the two dimensions and are comfortable with the interface, the main session begins. The main session consists of 39 trials; 36 original animations plus three repetitions. The order of the original animations is randomized for each participant to avoid order effects, and the first 3 animations are repeated at the end of the experiments in order to evaluate within-subject consistency. After the participants submit their ratings, the screen displays the confidence level Likert scale questions. When these are completed, the next trial begins. Between trials the robot resumes its initial neutral motionless standing posture, with eye LEDs switched off, so that participants can clearly perceive the onset and offset of the animation.

During the experiment, we also collect a few implicit measures: the number of times the participant replays each animation and the time elapsed from the start of the animation to the submission of the valence/arousal ratings, before the screen with the confidence level Likert scales is presented.

III. Analysis

Based on the data gathered, the following analyses are planned:

- Descriptive statistics of valence and arousal scores for each individual animation and for each of the nine classes of animations presented in Table I. These will be produced separately for women, men, and both groups.
- Interrater Reliability (IRR) statistics, testing the consistency between raters for valence and arousal. Additionally, we will test the consistency between men and women, as well as between the different combination of interactive modalities (motion, eye LEDs, sounds). Mean ratings derived from the IRR test will be the final set of annotations for the animations.
- Correlation between the implicit measures (reaction times and number of replays), the positive/negative affect scores, and the expressed levels of confidence. This will provide objective measures of reliability for the subjective ratings.

IV. Conclusions

Reflecting on the limitations of our work, we reckon one might find the choice of the two-dimensional model oversimplified, since it has been shown that subtle variations between certain emotions that share common core affect (e.g., fear, anger) might not be captured in less than four dimensions [7]. Although core affect is considered still valid as a low dimensional structure that represents a single component of the full-blown emotional episode [3], nonetheless, in our future work we would like to explore how additional dimensions, like dominance and unpredictability, could contribute in terms of discriminating information. Another limitation might be the fact that we did not systematize and balance the animation set in terms of the interactive modalities included in the animations, but we plan to address this issue in our analysis, where we will look more closely at the contribution of the individual modalities and their combinations.

Furthermore, in our future work we plan to study how emotional animations are perceived within a given emotion-inducing context and in this direction, the annotated set will allow us to test different settings, such as the impact of appropriate or inappropriate emotional expressions.

Acknowledgements

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 674868 (APRIL).

References

[1] C. Breazeal, Role of expressive behaviour for robots that learn from people, Philosophical Transactions of the Royal Society B: Biological Sciences, vol. 364, no. 1535, pp. 35273538, Dec. 2009.

[2] A. Beck, A. Hiolle, A. Mazel, and L. Cañamero, Interpretation of Emotional Body Language Displayed by Robots, in Proceedings of the 3rd International Workshop on Affective Interaction in Natural Environments, New York, NY, USA, 2010, pp. 3742.

[3] J. A. Russell and L. F. Barrett, Core affect, prototypical emotional episodes, and other things called emotion: dissecting the elephant, J Pers Soc Psychol, vol. 76, no. 5, pp. 805819, May 1999.

[4] J. Monceaux, J. Becker, C. Boudier, and A. Mazel, Demonstration: First Steps in Emotional Expression of the Humanoid Robot Nao, in Proceedings of the 2009 International Conference on Multimodal Interfaces, New York, NY, USA, 2009, pp. 235236.

[5] A. Betella and P. F. M. J. Verschure, The Affective Slider: A Digital Self-Assessment Scale for the Measurement of Human Emotions, PLOS ONE, vol. 11, no. 2, p. e0148037, Feb. 2016.

[6] D. Watson, L. A. Clark, and A. Tellegen, Development and validation of brief measures of positive and negative affect: the PANAS scales, J Pers Soc Psychol, vol. 54, no. 6, pp. 10631070, Jun. 1988.

[7] J. R. J. Fontaine, K. R. Scherer, E. B. Roesch, and P. C. Ellsworth, The world of emotions is not two-dimensional, Psychol Sci, vol. 18, no. 12, pp. 10501057, Dec. 2007.