PetRo's Multimodal Emotional Language

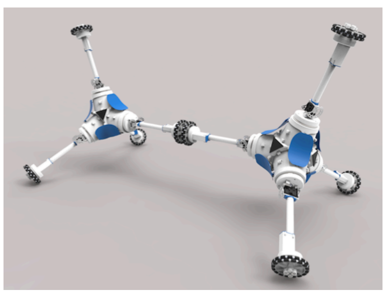
Ben Salem, *Member, IEEE,* Jonathon Stelling, Eliot Quinn, Tarik Safir and Kit McBride

***Abstract*— A robotic language of emotions could be as an effective mean to communicate a robot's internal state and condition. Such a communication would be beneficial within the context of a human robot interaction. We propose a vocabulary of emotions, and report on experiments about their recognition.**

I. Introduction

PetRo is a homogeneous modular **Pet**-Like **Ro**bot, with identical modules; it is designed to be handled without training [1]. A combination of body language and sounds are used by the robot to communicate its status and current task.

Figure 1. A rendering of PetRo in a two modules assembly (dog configuration).



II. Expressive Language

An expressive language based on body movements and sounds is proposed. It is a *multimodal social interaction* driven by the robot morphology and overall purpose (see 2]). It should include body movements, postures, orientation, colour, and sound [4]. In this perspective, body movements are particularly important, as they are as reliable as facial expressions for the recognition of emotions [5]. The effective combination of movements and sound should deliver recognisable expressions of extraverted emotions and introverted moods (as specified in [1]). Emotions, within the limitations of PetRo two modules assembly, can be expressed with body postures, walking stride, head direction, silences, heart beats' noises, breathing noises, sound outbursts, distance from user (see [6]). We believe the limitation of the assembly's abstract appearance is not an issue as it is not detrimental to the user's perception of a "partner" in a robot– it is actually the robot's tasks performance that matter [7].

We focus on the following emotions: Anger, Curiosity, Disgust, Fear, Happiness, and Sadness. We also include the following *moods*: Ready and Low Power. Apart from body language, sounds can also be used in combination with body movements to reinforce the expression of various emotions. Non-Linguistic Utterance and in general Sematic-Free Utterances have been proposed to "facilitate rich communication and expression" between a user and the robot (see [2], [3]). We have implemented breathing and heart beat sounds with variations to express different emotions.

III. Evaluation of the Language

To evaluate this language we conceived two of experiments to evaluate the language in terms of recognisability. To do so we developed a first version of the language that we evaluated with N=29 participants in a pilot experiment, and a second version that we evaluated with N= 25 participants.

A. Pilot Experiment

We collected the following results, comparing the mean score of emotions rendered by a dog in a video and those rendered by a robot in an animation, running a related samples t-test: Anger (Mean score for the Dog movie 6.759, Mean score for the robot animation 2.000. Mean difference 4.759, $t = 6.093$, $df = 28$, $p<0.001$), Curiosity (Mean score for the Dog movie 7.724, Mean score for the robot animation 5.138. Mean difference 2.586, $t = 3.042$, $df = 28$, $p=0.005$), Disgust (Mean score for the Dog movie 2.793, Mean score for the robot animation 0.793, Mean difference 2.000, $t = 2.681$, $df = 28$, $p=0.012$), Fear (Mean score for the Dog movie 1.345, Mean score for the robot animation 2.345. Mean difference -1.000, $t = 1.573$, $df = 28$, $p=0.127$), Happiness (Mean score for the Dog movie 5.034, Mean score for the robot animation 6.034. Mean difference -1.000, $t = 0.996$, $df = 28$, $p=0.328$), Sadness (Mean score for the Dog movie 6.655, Mean score for the robot animation 3.276. Mean difference -3.379, $t = 5.666$, $df = 28$, $p<0.001$)

We decided therefore to refine the rendering of emotions by the robot's model to ensure that the recognition of emotions is significantly better than when watching the video of the dog. We also decided to add sounds as an additional modality to express emotions.

B. Second Experiment

We have refined afterwards the language, added sound and ran a second experiment.

We have defined three test conditions that we refer to as: (1) Movie of a companion dog (see fig. 2), (2) Silent animations of a 3D model of PetRo version 3.09 in a dog

B.S. is with Northumbria University, Newcastle, NE1 8ST, UK (corresponding author phone: +44 7971 199 164; e-mail: mail@bsalem.info).

J.S. and E.Q. are undergraduate students at Northumbria University, Newcastle, NE1 8ST, UK.

T.S. is an undergraduate student at the University of Manchester, Manchester, M13 9PL, UK. Work was conducted during his summer internship at Northumbria University.

K.M. is with Whitecroft Lighting Ltd., Asthon, OL7 0AX, UK. Work was conducted during a summer project at Northumbria University.

Salem, B., Stelling, J., Quinn, E., Safir, T., McBride, K. (2016). PetRo's Multimodal Emotional Language, In: Proc. IEEE RO-MAN 2016, The 25th IEEE International Symposium on Robot and Human Interactive Communication, 932-933.

configuration (see fig. 3), and (3) Animations & Sounds, combining the animation clips and sound tracks to express emotions.

For the emotions Anger, Curiosity, Disgust, Fear, Happiness, and Sadness, we ran a One Factor Within Subject ANOVA Test. For the moods Ready and Low Power, we ran a Paired Samples T Test. We present herein the results of both the descriptive and analytical statistics.

Anger: Mean scores for Dog movie 6.000 (SD 3.7648), for Robot silent animation 3.250 (SD 4.1729), for Robot animation & sound 3.458 (SD 3.8446). The mean scores for the three clips differed significantly at the 5% level: $F(2, 46) = 4.57$; p=.015 Partial $\eta^2 = .166$, which is a large effect.

Curiosity: Mean scores for Dog movie 6.520 (SD 3.5369), for Robot silent animation 6.480 (SD 3.4775), for Robot animation & sound 4.520 (SD 4.1041). The mean scores for the three clips did not differ significantly at the 5% level: $F(2, 48) = 2.686$; p=.078 Partial $\eta^2 = .101$, which is a medium effect.

Disgust: Mean scores for Dog movie 1.680 (SD 2.3224), for Robot silent animation 0.800 (SD 1.5546), for Robot animation & sound 0.320 (SD 0.9000). The mean scores for the three clips differed significantly at the 5% level: $F(2, 48) = 4.57$; p=.016 Partial $\eta^2 = .159$, which is a large effect.

Fear: Mean scores for Dog movie 0.720 (SD 2.1315), for Robot silent animation 2.720 (SD 3.8570), for Robot animation & sound 6.520 (SD 3.6982). The mean scores for the three clips differed highly significantly at the 1% level: $F(2, 48) = 23.963$; p=.000 Partial $\eta^2 = .500$, which is a large effect.

Happiness: Mean scores for Dog movie 4.920 (SD 3.6391), for Robot silent animation 7.840 (SD 2.8965), for Robot animation & sound 5.800 (SD 3.7749). The mean scores for the three clips differed highly significantly at the 1% level: $F(2, 48) = 5.166$; p=.009 Partial $\eta^2 = .177$, which is a large effect.

Sadness: Mean scores for Dog movie 7.667 (SD 3.1161), for Robot silent animation 4.000 (SD 3.6475), for Robot animation & sound 3.417 (SD 3.6226). The mean scores for the three clips differed highly significantly at the 1% level: $F(2, 46) = 11.629$; p=.000 Partial $\eta^2 = .336$, which is a large effect.

Ready: Mean scores for Robot silent animation 3.000 (SD(4.1028), for Robot animation & sound 4.520 (SD 4.5563). The mean scores for the animation & sound clip was greater than the mean score for the silent animation clip. Unfortunately, a related-samples *t* test showed no significance at the 5% level: t (24) = -1.570; p=.129 (two-tailed).

Low Power: Mean scores for Robot silent animation 3.920 (SD 3.2005), for Robot animation & sound 9.560 (SD 1.2936). The mean scores for the "Animation and Sound" clip for the internal state "No Power" was greater than the mean score for the "Silent Animation" clip. A related-samples *t* test showed high significance beyond the 1% level: t (24) = -9.494; p=.000 (two-tailed). The 95% confidence interval on the difference was [-6.8661, -4.4139], which does not include the value of zero specified by the null hypothesis.

Some of the participants wrote further comments in the space allowed to do so in the questionnaires – in a similar fashion to the pilot experiment. In the case of the Dog movie clips, 40.54% the comments related to the description of emotions perceived, and 64.86 to the description of body movements. In the animation clips of the robot, 7.23% related to emotions and 95.12% to body movements. In the animation and sound clips the results were: 11.76%, 67.65%, and 32.35% for sound.

IV. Discussion

The following emotions have been successfully recognised: Fear and Happiness. They score significantly better than the dog's expressions. However, Anger, Sadness are significantly better expressed by the dog. For Anger, it is possible to explain it by the saccadic movements forward of the head and shoulders that is difficult to reproduce in the robot. As for Sadness, the expression of the emotion in the dog is emphasised by the adopted stride that is clearly visible during the movie. One emotion is clearly problematic: Disgust – it is poorly recognised, with either the dog or the robot. One initial explanation is that this is a triggered emotion, one that requires an object that is the focus/target of the emotion. It is mostly confused with Curiosity (28%, 16% and 16% in each of the test conditions). In the case of the movie clip, initially the dog moves towards the object to check it out before expressing disgust at it, which is somewhat misleading. Further work is necessary in improving the rendering of emotions, the soundtracks used and in evaluating if dog ownership is a factor in emotion recognition. We would also like to evaluate the level of confidence participants have in recognising emotions.

Acknowledgment

We would like to thank Harry CRAWFORD for his support during the completion of version 3.09 of PetRo.

References

[1] B. Salem, "PetRo: development of a modular pet robot", in *Proc. 23rd IEEE International Symposium on Robot and Human Interactive Communication,* Edinburgh, UK, 2014, pp.483-488.

[2] R. Read, and T. Belpaeme, "People interpret robotic non-linguistic utterances categorically", in *Proc. 8th HRI,* Tokyo, JP, 2013, pp. 209-210.

[3] S. Yilmazyildiz, R. Read, T. Belpeame, and W. Verhelst, "Review of semantic-free utterances in social Human-robot interaction", *Int. Journal of Human-Computer Interaction,* vol. 32, pp. 63-85, 2016.

[4] C.L. Bethel, R.R. Murphy, "Survey of non-facial/non-verbal affective expressions for appearance-constrained robots", *IEEE Transactions on Systems, Man, and Cybernetics —Part C: Applications and Reviews*, vol.38, no. 1, pp.83-92, January 2008.

[5] B. de Gelder, "Why bodies? Twelve reasons for including bodily expressions in affective neuroscience", *Phil. Trans. R. Soc. B*, vol. 364, pp. 3475-3484, 2009.

[6] A. Vinciarelli, M. Pantic, and H. Bourlard, "Social signal processing: Survey of an emerging domain", *Image and Vision Computing,* vol. 27, pp. 1743-1759, 2009.

[7] R.A. Paauwe, J.F. Hoorn, E.A. Konijn, and D.V. Keyson, "Designing robot embodiments for social interaction: Affordances topple realism and aesthetics", International Journal of Social Robotics, vol. 7, pp. 697-708, 2015.