Generation of Meaningful Robot Expressions with Active Learning

Giovanni Saponaro
gsaponaro@isr.ist.utl.pt

Alexandre Bernardino
alex@isr.ist.utl.pt

Institute for Systems and Robotics, Instituto Superior Técnico
Av. Rovisco Pais, 1. 1049-001 Lisbon, Portugal

ABSTRACT

We propose a mechanism to communicate emotions to humans by using head, torso and arm movements of a humanoid robot, without exploiting its facial features. To this end, we build a library of pre-programmed robot movements and we ask people to attribute emotional scores to these initial movements. The answers are then used to fine-tune motion parameters with an active learning approach.

Categories and Subject Descriptors

H.1.2 [**Models and Principles**]: User/Machine Systems—*Human information processing*; I.2.6 [**Artificial Intelligence**]: Learning—*Parameter learning*

General Terms

Human Factors, Performance

Keywords

Active learning, emotions, robot movements

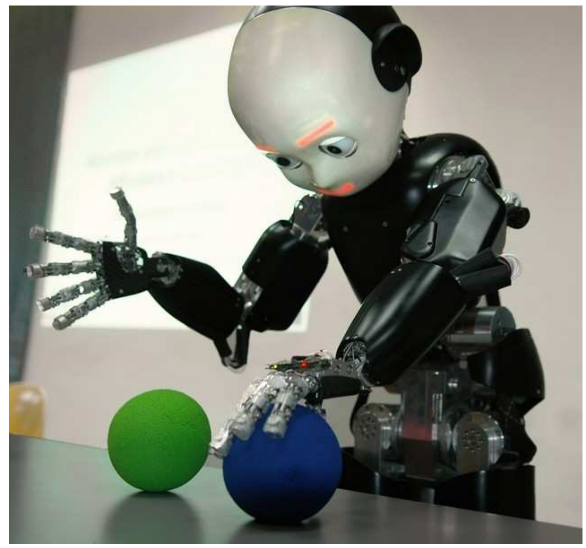

Figure 1: The iCub (picture by Lorenzo Natale).

1. INTRODUCTION

The iCub (Fig. 1) is an open source, child-like humanoid robot endowed with 53 degrees of freedom including fully articulated eyes and head, and the ability to perform dexterous manipulation [3] as well as articulatory gestures.

So far, the iCub has been used to display some basic emotions [1], however this task has been mainly concerned with the robot *face* by controlling its eyebrow LEDs, mouth LEDs and eyelid servomotors, all features that are indeed extremely informative for the transmission of feelings. With this work, by contrast, we intend to explore the emotion capabilities of *joint movements* located in the head, neck, torso and arms of the robot, and how well human users interpret these movements when face expressions are disabled.

We design a library of pre-defined actions and survey a number of people about what they feel when the robot performs them. Subsequently, we feed the human opinions back into an active machine learning program that chooses which actions require more corrections and information, in order to fine-tune action movement parameters (joint positions, joint velocities, timings) and ultimately convey the intended emotions more clearly.

HRI'11, March 6–9, 2011, Lausanne, Switzerland.
ACM 978-1-4503-0561-7/11/03.

In this report, we outline the conceptual blocks of the work.

2. CONVEYING EMOTIONS WITH MOVEMENTS

Because we are interested in mapping simple robot actions to emotions that we wish to transmit successfully, the first crucial aspect of the work consists in designing a library of basic motions and expected perceived emotions, displayed in Table 1. Interestingly, some of the gestures have a similar or overlapping meaning, such as "V sign" and "fist pump". In these cases, we also question: what is the most effective way for the robot to transmit a certain emotion (e.g., "success") without using face features? Which of the two candidate actions has more room for improvement, given the mechanical and kinematic capabilities of the robot?

As we make the robot perform the actions and we ask human users to respond to a questionnaire about the perceived emotions, we are implicitly studying two issues that must be kept in mind when dealing with robot gestures: (i) whether simple robot actions, in the absence of facial cues, convey the expected emotions to viewers; (ii) what contextual and motion characteristics aid gesture understanding the most [2].

To survey human interpretation of the robot movements, we present people with a five-level Likert questionnaire. The

Table 1: Library of robot action movements

Action	Description	Expected perceived emotion(s)
nod	head tilts up and down	agreement, acknowledgement
punch	rapidly extend fist in front of robot	anger, unease, fear
look out	abruptly deviate robot head and gaze to a side	disregard, distraction
thumbs up	show fist and move thumb up	approval
thumbs down	show fist and move thumb down	disapproval
V sign	raise arm and show victory sign	success, happiness
fist pump	raise fist in front of torso then drop it down fast	celebration, success, victory

Table 2: Parameters of the "nod" action

Parameter	Meaning
$x_0^{(0)}$	initial position of first neck joint
$x_0^{(1)}$	final position of first neck joint
$\dot{x}_0^{(0)}$	initial velocity of first neck joint
$t_{0\to 1}$	time to transition from (0) to (1)
$\dot{x}_0^{(1)}$	final velocity of first neck joint
$t_{1\to 0}$	time to transition from (1) to (0)

robot displays actions selected from the library of Table 1 and we ask people to rate their level of agreement to a number of statements relative to emotions: "This action expresses anger", "This action expresses robot interest in the person", "This action expresses success", etc.

3. LEARNING MOTION PARAMETERS

The scores that result from the human survey described above can be employed in an *active* Bayesian Network learning program. In the active learning framework [4], the learner has the ability to guide the instances it gets, by querying for a particular input rather than proceeding randomly or sequentially from a set.

We model simple Bayesian Networks, each one composed by two nodes only: A (action) $\rightarrow$ E (emotion). Node A includes the type of action and its parameter values, as follows: it contains the specific type of action $T = t_i$, where $i = 1 \dots M$ (M: number of possible actions), as in Table 1, and its parameter values $V_{ij} = v_{ijk}, j = 1 \dots P_i$ (P_i: parameter index for action i, as in Table 2), $k = 1 \dots K_{ij}$ (K_{ij}: discretized index of parameter j for action i). We use a unique index n to count all the possible action–parameter nodes A. Node E encodes a pre-defined set of possible emotions e_l, $l = 1 \dots L$, similar to the last column of Table 1.

The probability distribution $P(E|A)$ is modelled as a multinomial distribution $P(E = e_l|A = a_n) = \theta_{ln}$, where θ are the parameters, l is the emotion index, n is the action–parameter index and $\sum_l \theta_{ln} = 1$ for each action a_n.

The active learning algorithm selects the action–parameter a_n such that the expected entropy gain is maximum. The entropy gain is the difference between current entropy (i.e., before applying the learning step), and expected posterior entropy after a trial. The update rule is:

$$P(E = e_l|A = a_n) \leftarrow \frac{P(E = e_l|A = a_n) \cdot \#a_n + s}{\#a_n + 1}$$

where $\#a_n$ is the number of trials performed with $A = a_n$ up to this point and s is the Likert score resulting from the current trial answer (normalized to a probability).

An advantage of the proposed approach is that we do not have to survey *all* people about about *all* the possible action–emotion networks. Instead, the active learning module (i) selects an action–emotion network that requires improvement; (ii) selects a particular parameter to fine-tune for that action; (iii) updates the parameter value. For example, the program could (i) choose to improve the effectiveness of the "punch" action; (ii) among the parameters of this action, select the elbow velocity parameter; (iii) increase or decrease said parameter by using the Likert score from the human user feedback.

4. CONCLUSIONS

We address the problem of communicating emotions with a humanoid robot *without using the facial features* but employing movements of head, arms and torso. The proposed method is summarized as follows: (i) design a library of simple robot movement actions corresponding to a ground truth of emotion(s); (ii) survey a number of people with a Likert questionnaire about how they interpret the movements in the library; (iii) according to the survey scores, improve the robot trajectories (positions, velocities and timings) with an active learning algorithm that selects the next action to display from the library, and the parameter to tune.

5. ACKNOWLEDGMENTS

The authors thank João Sequeira for the helpful contribution. This work was supported by Fundação para a Ciência e a Tecnologia (ISR/IST pluriannual funding and grant SFRH/BD/61910/2009).

6. REFERENCES

[1] P. Ekman et al. *Emotion in the Human Face: Guidelines for Research and an Integration of Findings.* Pergamon Press, New York, 1972.

[2] J. Li and M. Chignell. Communication of Emotion in Social Robots through Simple Head and Arm Movements. *International Journal of Social Robotics*, 2010.

[3] G. Metta et al. The *iCub* Humanoid Robot: An Open-Systems Platform for Research in Cognitive Development. *Neural Networks*, 23(8-9):1125–1134, 2010.

[4] S. Tong and D. Koller. Active Learning for Parameter Estimation in Bayesian Networks. In T. K. Leen, T. G. Dietterich, and V. Tresp, editors, *Neural Information Processing Systems (NIPS)*, pages 647–653. MIT Press, 2000.