Generation of Meaningful Robot Expressions with Active Learning

Giovanni Saponaro, Alexandre Bernardino
VisLab, Institute for Systems and Robotics, Instituto Superior Técnico, Lisbon, Portugal

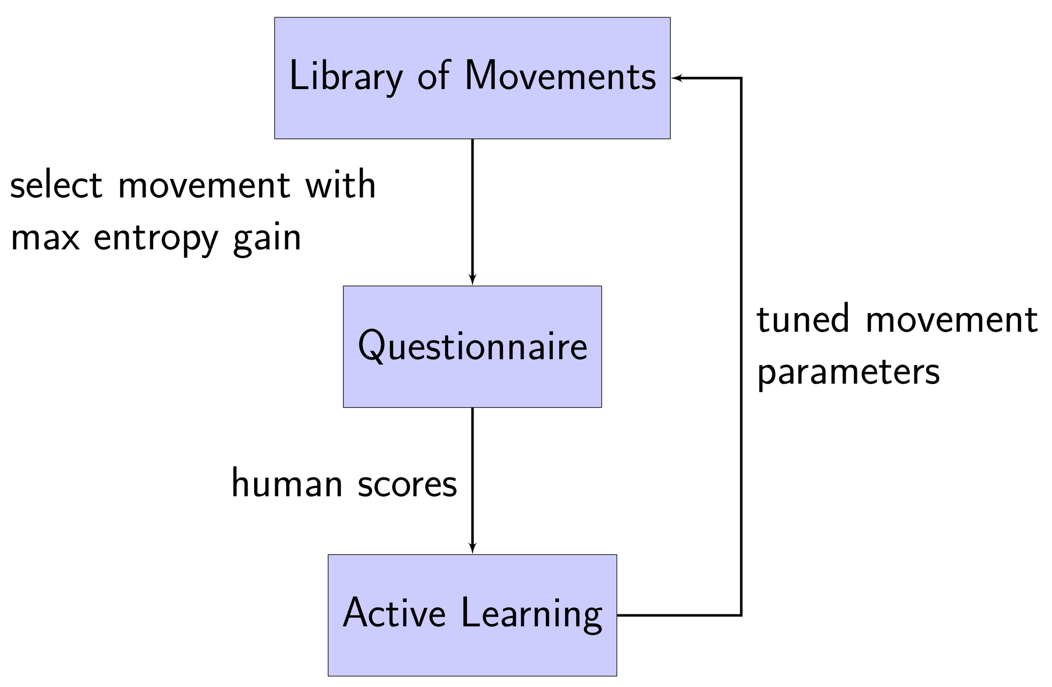
Overview

We address the problem of communicating emotions with a humanoid robot merely with its body joint movements, without facial information, querying human users for emotional scores attributed to the movements. Machine learning can help convey the intended emotions more clearly, by selecting the next actions and parameters that need tuning and rewarding successful action–emotion matches.

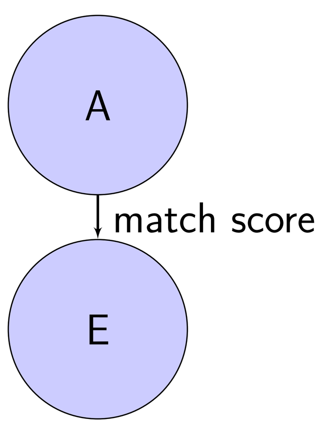
Motivation

- Transmit emotions to human observers with the iCub humanoid robot
- Explore expressivity (or lack thereof) of body movements when facial expressions are disabled

Setup and Proposed Approach



Each combination is a simple Bayesian Network with an Action and an Emotion node:



- Design a library of pre-defined movements

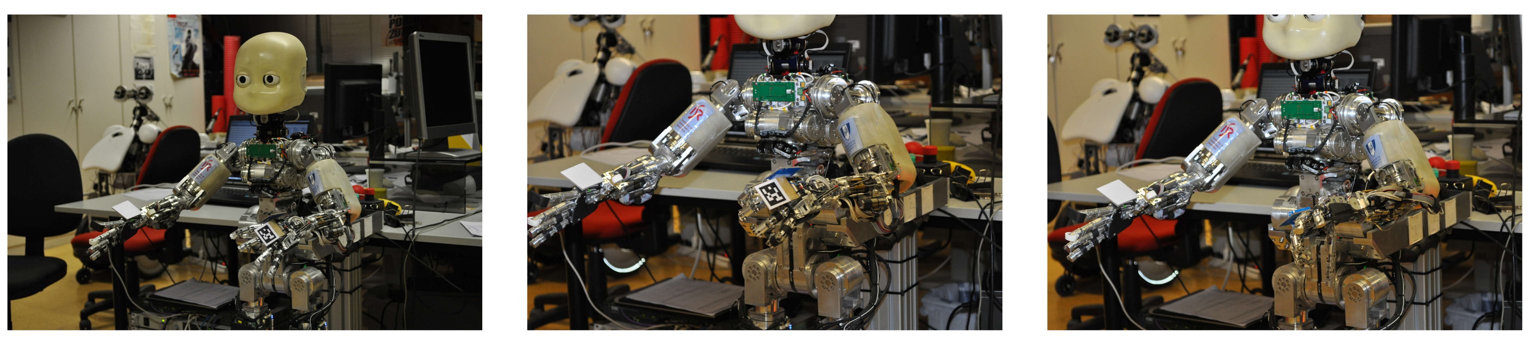


Figure: The iCub performing the "thumbs down" gesture.

- Survey human users about what they feel when the robot performs said movements

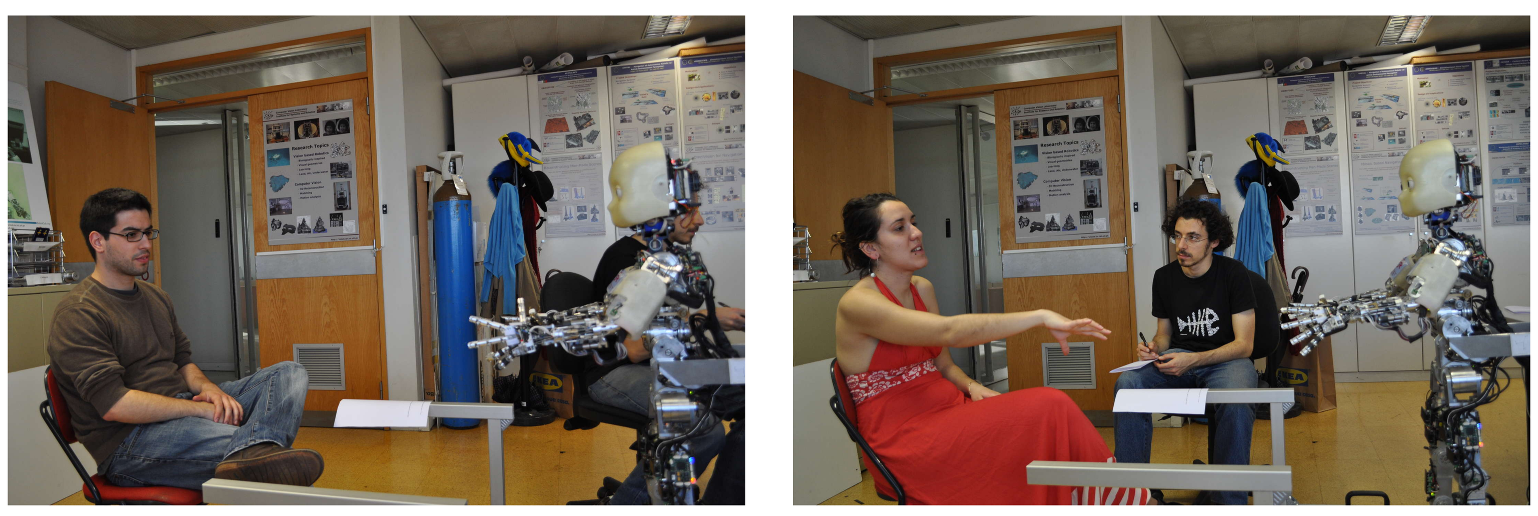


Figure: Human users being interviewed about what emotion they perceive from robot movements.

- Input human opinions into an active machine learning program that selects the next action to show (the one requiring more correction, i.e., with most entropy)
- Active learner can also choose an action movement parameter and a discretized value to fine-tune

Example Library of Movements and Parameters

Table: Library of robot action movements

Action a_n	Description	Expected perceived emotion e_l
nod	head tilts up and down	agreement
punch	rapidly extend fist in front of robot	anger
look out	abruptly deviate robot head and gaze to a side	distraction
thumbs up	show fist and move thumb up	success
thumbs down	show fist and move thumb down	disapproval

Table: 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)

User Survey

- Human subjects see robot actions and respond to a Likert questionnaire
- After each robot movement, a user is asked to evaluate statements such as "This action expresses anger" with a score among [strongly disagree] [disagree] [neither agree nor disagree] [agree] [strongly agree]
- Responses are sent as a probability vector to the Active Learning module, which updates action–emotion parameters
- We interviewed 20 people (non-roboticists): 50% males, 50% females, mean age 29.7, stddev 4.49
- Type of study: within-subject with self-assessment evaluation

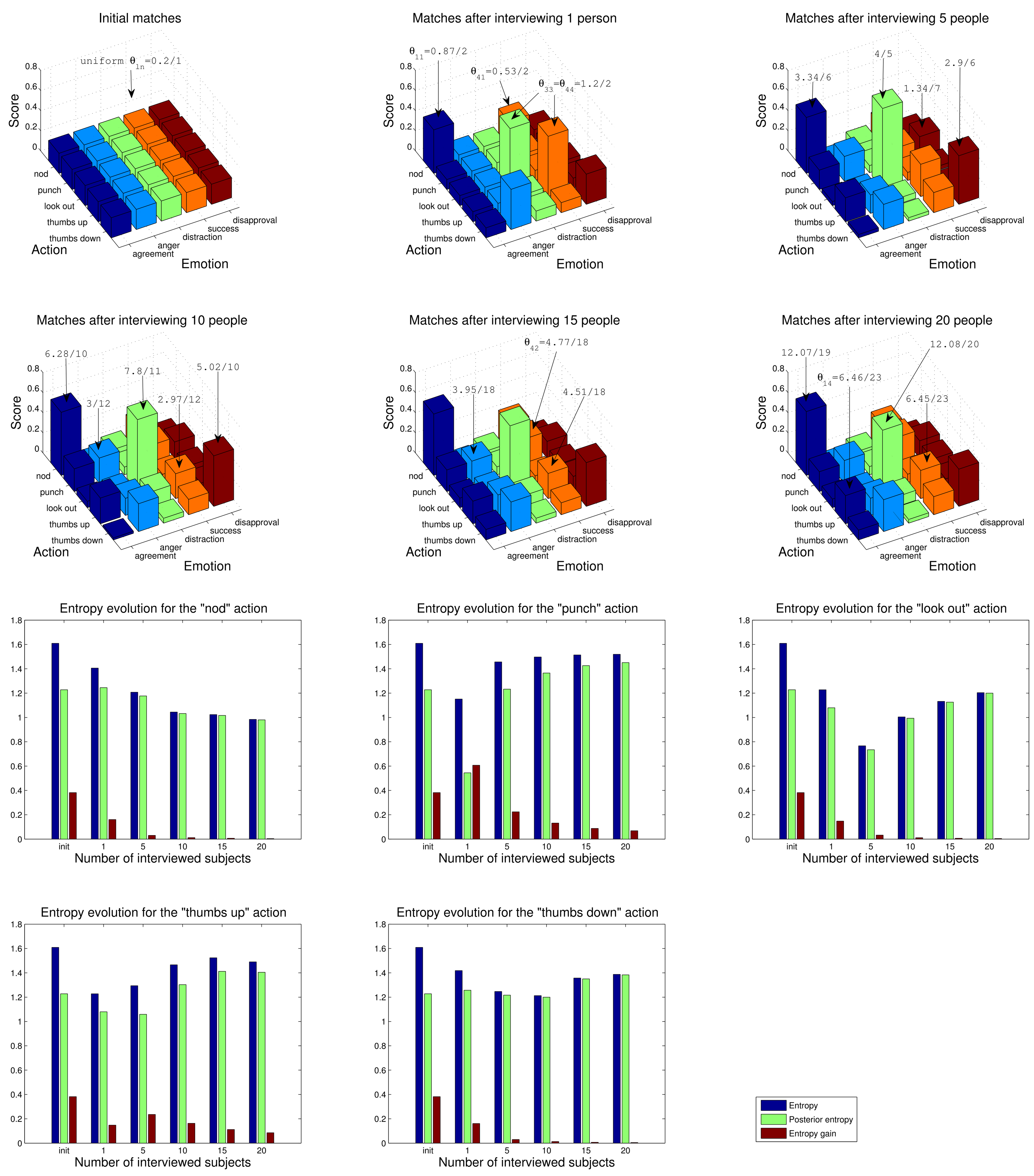
Active Learning

$$P(E=e_l|A=a_n) = [\boldsymbol{\theta}_{ln}] = \begin{bmatrix} \theta_{11} & \cdots & \theta_{L1} \\ \theta_{12} & \cdots & \theta_{L2} \\ \vdots & \ddots & \vdots \\ \theta_{1A} & \cdots & \theta_{LA} \end{bmatrix}$$

- Action–emotion matches modelled as a multinomial distribution
- Rows: actions (with a discretized joint parameter), columns: emotions, $\Sigma_l \theta_{ln} = 1 \quad \forall$ action
- *Active*: learner selects the row yielding highest entropy gain as next query, we update the parameters of that row
- $\frac{P(E=e_l|A=a_n)}{\#a_n} \leftarrow \frac{P(E=e_l|A=a_n)+s}{\#a_n+1}$: update step, 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)
- The framework is scalable to use a matrix with hundreds of rows, where clusters of rows encode the same action with different joint parameters, and can be awarded/penalized

Results

- Head actions are associated to emotions clearly, arm gestures are often ambiguous
- Unclear actions are chosen and displayed more times than clear ones, as reflected in the higher denominators of match scores



Future Work

- Improve corpus of robot movements, explore using other limbs (legs), systematically analyze the human affective experience with or without robot facial expressions
- Study the degree of robot emotional expressivity due to its appearance vs the mechanical range of permitted movements
- Try different parameter initialization: uniform vs expert knowledge
- Explore optimization strategies other than maximum-entropy criterion: artificial neural networks, reinforcement learning

Summary

We model a mapping between robot movements (with no facial expressions) and emotions perceived by human users. Match parameters are updated with a questionnaire, and as correspondences take shape we can observe (i) which body joint actions are expressive, (ii) which actions should be performed by a robot in order to transmit a given emotion.
As opposed to random action selection or querying for a fixed sequence of actions, the Active Learning framework gives the system the ability to inquire about movements that are ambiguous, showing them more often than easily-perceived ones.

References

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

[2] 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.

This work was supported by Fundação para a Ciência e a Tecnologia (ISR/IST pluriannual funding and grant SFRH/BD/61910/2009).