iCub: Learning Emotion Expressions using Human Reward

Nikhil Churamani, Francisco Cruz, Sascha Griffiths and Pablo Barros

Abstract— The purpose of the present study is to learn emotion expression representations for artificial agents using reward shaping mechanisms. The approach takes inspiration from the TAMER framework for training a Multilayer Perceptron (MLP) to learn to express different emotions on the iCub robot in a human-robot interaction scenario. The robot uses a combination of a Convolutional Neural Network (CNN) and a Self-organising Map (SOM) to recognise an emotion and then learns to express the same using the MLP. The objective is to teach a robot to respond adequately to the user's perception of emotions and learn how to express different emotions.

I. Introduction

With the advances in human-robot interaction research, artificial agents are slowly but surely becoming an integral part of human life. We are going to interact with these agents on a daily basis, in one form or the other. As our interaction and thus inadvertently, our dependency on agents increases, these agents need to blend into the social environments surrounding them, as naturally as possible. Thus, agents need to be able to factor in emotions and sentiments while dealing with their human-centred environment so as to make well-informed decisions.

This study makes use of the iCub robot head and its on-board capabilities to encode different emotional expressions for seven "Universal Emotions", as proposed by Ekman [1] viz. *Anger*, *Disgust*, *Fear*, *Happiness*, *Sadness* and *Surprise* along with *Neutral* as an additional emotional state to accommodate for the absence of any emotion.

Also, this work explores Reinforcement Learning (RL) strategies, particularly in the context of *reward shaping* [2] approaches to train an MLP to learn different emotion expression states for perceived emotions. The agent makes use of the on-board camera to capture an image of the face of the user and uses the CNN, to come up with a feature representation for the face image. These feature representations are then used to train the SOM, as presented in [3], forming clusters representing each of the aforementioned seven emotions. The MLP is then used to come up with the most appropriate expression representation related to an action to be used for the emotion. This work builds on the work by Barros et. al [3] and uses the underlying CNN and SOM model to enhance the proposed system.

The implementation is inspired by the TAMER [4]–[6] algorithm for modelling a "*learning by human reward*" approach so as to train the MLP.

The authors are with the Knowledge Technology Institute, Department of Informatics, University of Hamburg, Hamburg, Germany. {5churama, cruz, griffiths, barros} @informatik.uni-hamburg.de
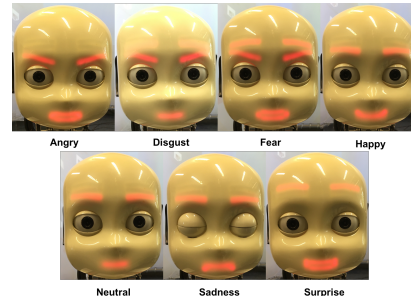

Fig. 1: iCub Robot Head encoding emotions

II. Emotion Expression using icub

In the context of this study, we make use of the iCub robot head [7] which uses multiple underlying electronic control boards corresponding to LEDs representing the left and the right eyebrows and the mouth. The iCub robot head uses an *emotionInterface* module which is able to map different emotion states to expression representations via the LEDs available on the face. Different subsystems like the mouth, eye-lids, left eyebrow, right eyebrow or the whole system can be used alternatively to encode different emotion states by setting LED flags for these subsystems. The encodings for emotion expressions used in this study are shown in Fig.1.

The study assumes that there is a unique response, taking the form of an "expression selection", to each input emotion state and thus, one particular facial representation shall uniquely identify a particular emotion. Forming a particular facial representation (using a combination of lights, selecting them one by one) can also be broken down into subtasks which could be learned as well. In this study, a particular combination of lights is selected as the lowest abstraction level for action selection for the sake of simplicity.

III. Proposed Model

The motivation behind this study is to explore the possibility of training artificial agents to perform actions, while interacting with a human user. In the context of emotions, different people express and perceive emotions differently and thus, the agent needs to adapt to this variance and customise its learning depending on the user.

The iCub captures an image of the user's expression which is fed to a pre-trained CNN giving it a feature vector representation. The face feature representations are then fed to the SOM where it is expected that clusters emerge, each pertaining to a particular emotion. User interactions are modelled by taking the Best Matching Unit (BMU) from the SOM for training an MLP to predict the action which best
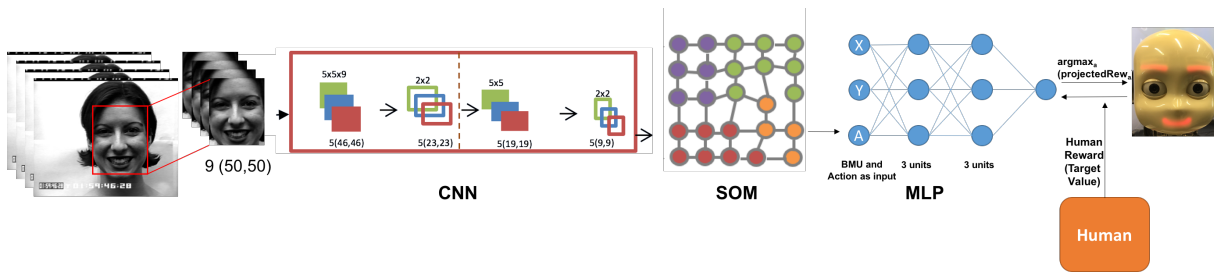
Fig. 2: Complete Model with CNN face representation, SOM clusters and the MLP predicting a reward and selecting an action learning from human reward
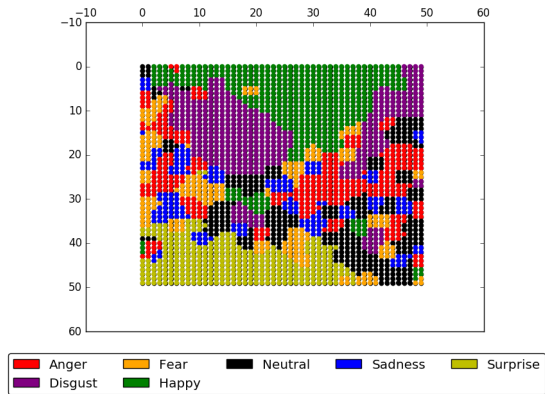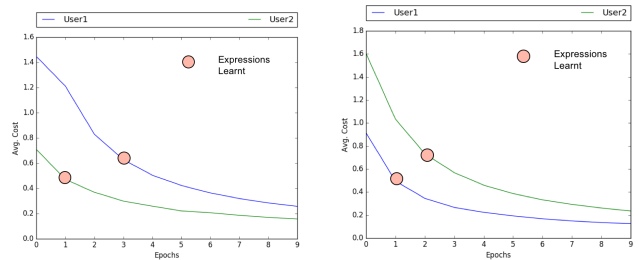

Fig. 3: SOM Cluster Map

(a) Users with no prior knowledge about the system

(b) Users with prior knowledge about the system

Fig. 4: Preliminary Results: Each epoch corresponding to 100 interactions for calculating the Avg. cost

mimics the user's expression. The user rewards the robot's action and the MLP is trained to select the correct expression by learning to predict this reward. The CNN is trained using the Cohn-Kanade dataset [8], in this case consisting of approximately 1000 different images. The SOM on the other hand is trained each time for a specific user.

The CNN consists of two convolution layers (with max-pooling) also making use of L1 and L2 normalisation for each layer respectively. The resulting feature vector representation is used to train the SOM, forming clusters for different emotions (see Fig. 3). For each user interaction, the BMU is computed (position) and fed to the MLP along with a possible action (expression representation) resulting in a predicted reward value ($r \; \epsilon \; [-2, 2]$). The agent then tries all possible actions and chooses the action with the highest reward value and performs that action. Fig. 2 depicts the complete model.

Once the iCub performs an action, the user is expected to reward it, thus giving it a target value to achieve. This is done by asking the user to mimic the robot giving it information about how much the action performed differs from the intended action. The iCub again captures the user mimicking the agent and evaluates the BMU for the reward. The reward value ($r \; \epsilon \; [-2, 2]$) is computed by using the *Normalised Euclidean Distance* between the BMUs and thresholding it i.e. the lesser the distance between the BMUs, the more positive the reward. The aim of the MLP learning algorithm is thus, to approximate an optimal reward function which is able to estimate the user's reward for each emotion-expression pair.

IV. PRELIMINARY RESULTS

Some preliminary studies were performed with users for the model providing positive and motivating results (see Fig. 4). Although the results were promising and significantly reduced the time needed for training, the method still required more than 100 interactions per user to learn meaningful expressions. This number is expected to decrease with improvements in training methodologies and by collecting more data for training.

REFERENCES

[1] P. Ekman, "An argument for basic emotions," *Cognition & emotion*, vol. 6, no. 3-4, pp. 169–200, 1992.
[2] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT press, 1998.
[3] P. Barros and S. Wermter, "Developing crossmodal expression recognition based on a deep neural model," in *Adaptive Behaviour*, (to appear), 2016.
[4] W. B. Knox and P. Stone, "Tamer: Training an agent manually via evaluative reinforcement," in *2008 7th IEEE International Conference on Development and Learning*, pp. 292–297, IEEE, 2008.
[5] W. B. Knox, "Tamer: Training an agent manually via evaluative reinforcement," in *Learning from Human-Generated Reward*, Ph.D Thesis, 2012.
[6] W. B. Knox, P. Stone, and C. Breazeal, "Training a robot via human feedback: A case study," in *Social Robotics*, pp. 460–470, Springer, 2013.
[7] iCub.org, "iCub open source cognitive humanoid robotic platform." https://www.wiki.icub.org, Accessed 2016-06-03.
[8] P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar, and I. Matthews, "The extended cohn-kanade dataset (ck+): A complete dataset for action unit and emotion-specified expression," in *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition-Workshops*, pp. 94–101, IEEE, 2010.