Towards Learning Affective Body Gesture

Andrea Kleinsmith
University of Aizu, Database Systems Lab
Aizu Wakamatsu, 965-8582 Japan
m5061202@u-aizu.ac.jp

Nadia Bianchi-Berthouze
University of Aizu, Database Systems Lab
Aizu Wakamatsu, 965-8582 Japan
nadia@u-aizu.ac.jp

Abstract

Robots are assuming an increasingly important role in our society. They now become pets and help support children healing. In other words, they are now trying to entertain an active and affective communication with human agents. However, up to now, such systems have primarily relied on the human agents' ability to empathize with the system. Changes in the behavior of the system could therefore result in changes of mood or behavior in the human partner. This paper describes experiments we carried out to study the importance of body language in affective communication. The results of the experiments led us to develop a system that can incrementally learns to recognize affective messages conveyed by body postures.

1. Introduction

Many researchers have been trying to examine the relation between emotion and nonverbal cues. In the same vein of work as (Breazeal and Scassellati, 1999), researchers are explicitly exploiting the empathy of human caretakers for the human-like characteristics of the system and are trying to identify the causes for this empathy. This is achieved by codifying expressions to represent some emotional state. The robots often do not have any real learning capability, besides hard-wired evolution. These studies also investigate the effects of changes in the system on the behavior of the human caretaker. In those systems, affective behaviors are hence mainly passive, and the system's ability to react to its human partner's emotion is generally very limited or missing, resulting in a social interaction that fails over time. Indeed, over the long run, habituation takes place – the human partner gets bored – because the system does not seem to react to the user, or only in clearly stereotypical ways. In that sense, current systems miss the bi-directionality and/or unpredictability typical of human social interaction (Bianchi-Berthouze and Lisetti, 2002).

Suggested by (Picard, 1997), a different line of research exploits information from physiological cues to detect and react to emotions. For example pulse, galvanic skin response, temperature, and blood pressure can be measured by sensors to understand changes in the affective state of the human.

Meanwhile, little attention has been placed on the visual signals to be extracted from body posture. Dance and choreographic studies have shown that it is a powerful and frequently used means of communication.

2. Movement Features and their Affective Messages

An interesting aspect of body language is that it does not take body structural similarities between agents for successful interaction to occur. Humans can read affective body language in bodies that are not human-like. Accordingly, the modeling of affective body language in robots is very interesting because it requires going beyond pure recognition of the posture to explore a more general mapping. Studies, on dancing motions (von Laban, 1988) in particular, have shown that a movement conveys a different affective messages when its features, e.g. its spatial dimensions, are modified. Other studies, such as described in (Nakata, 2002), have shown that the same motion was accepted with different degrees of naturalness if its features, such as speed, were changed. These studies have been performed on non-anthropomorphic bodies, demonstrating the extent of humans' abilities to empathize to non-human systems.

In this study, we carried out preliminary experiments to understand how humans recognize affective postures. Specifically, we aimed at identifying the features of body language that lead an observer to associate an emotional state to a given posture or motion. To this end, we created an avatar with a human-like body. Facial expressions were not used so as to focus only on body motion. The avatar was animated using key-framing and inverse kinematics.

Table 1 summarizes the observations made by seven Japanese students on 5 types of affective animation. From the users' observations, we identified as relevant the following motion features: Complexity (C), Irregularity (I), Symmetry (S), Speed (Sp),

Labels	C	I	S	Sp	A	R
angry	↗	↗	↘	↗	–	–
selfish	↗	↗	↘	–	–	↗
happy	↗	↗	↘	–	↗	↗
scared	–	–	–	↗	↗	–
sad	↘	↘	–	↘	↗	↗

Table 1: The table summarizes subjects' observations. (↗): necessary, (–): irrelevant, (↘): not desired.

Amplitude (A), Rigidity (R). The complexity feature refers to the number of body parts that were involved in the motion. It was observed that, in some cases, the simplification of the body movement was making the recognition of the affective state more difficult. The speed feature was pointed out as important in the case of *angry* and *sad* animations. In the first case, a slow tapping of the foot was perceived as a rhythmical accompanying movement. Symmetry and irregularity were shown to be very important features both for the naturalness of the gesture and for effectively conveying an affective state. In these cases, the movement was repeated continuously without any variation. Rigidity not only resulted in the motion being effectively seen as unnatural but also decreased the intelligibleness of the affective state.

3. Capturing Emotional States

We implemented a system that incrementally learns the mapping function between body postures and emotional labels. A video camera captures a human made expressive posture and sends it to the computational system. The system analyzes and creates a posture signature that describes the relative position of the body joints. Details about the posture description are in (Bianchi-Berthouze et al., 2003). However, the algorithm proposed is very simple, and is not the focus of our work at this stage. The signature is hence mapped into an emotional label. The recognized emotion can be further used, according to the motivation or goal of the system, to select an appropriate reaction, e.g. to play a specific piece of music to change the user's state.

We see the mapping of posture features into emotional labels as a categorization problem. We propose to use a modified version (Berthouze and Tijsseling, 2002) of Categorizing and Learning modules (CALM) (Murre et al., 1992), that can self-organize inputs into categories. A CALM network consists of several CALM modules. While the topology of a CALM architecture is fixed, connections between modules are learnt. To improve its performance, the system can receive two types of feedback from its user. The first type of feedback is verbal feedback and is sent by the user directly to the system to explicitly indicate the correct emotional label when it outputs the wrong label. The second type of feedback is postural feedback and corresponds to changes in the user body postures as a reaction to the system's actions.

4. Conclusions

We tested the Mood Mapping module with 108 different postural images of an agent with marked joints involving 3 types of affective postures: *happy*, *angry* and *sad* postures. Only 1 error occurred: the posture was related to a *sad* emotion but was classified as *angry* by the model. The final aim of this work would see the system as an active actor in the interaction with the human. To increase the complexity of the possible postures to be modeled, we are now beginning to exploit motion features by analyzing data acquired with both a 3D motion capture system and video. This will require us to deal with the temporal dimension of the motion features already detected as relevant in our preliminary experiments.

References

Berthouze, L. and Tijsseling, A. (2002). Acquiring ontological categories through interaction. *3DForum*, 16(4):141–147.

Bianchi-Berthouze, N., Fushimi, T., Hasegawa, M., Kleinsmith, A., Takenata, H., and Berthouze, L. (2003). Learning to recognize affective body postures. *Proc. of the IEEE Intl. Symposium on Computational Intelligence for Measurement Systems and Applications, Lugano, Switzerland.*

Bianchi-Berthouze, N. and Lisetti, C. (2002). Modeling multimodal expression of users' affective subjective experience. *International Journal on User Modeling and User-Adapted Interaction*, 12(1):49–84.

Breazeal, C. and Scassellati, B. (1999). How to build robots that make friends and influence people. *International Conference on Intelligent Robots and Systems,Kyonjiu, Korea.*

Murre, J., Phaf, R., and Wolters, G. (1992). Calm: Categorizing and learning module. *Neural Networks*, 5:55–82.

Nakata, T. (2002). Generation of whole-body expressive movement based on somatical theories. *Proceedings of the 2nd international workshop on epigenetic robotics*, pages 105–114.

Picard, R. (1997). *Affective Computing.* MIT Press.

von Laban, R. (1988). *The mastery of movement.* Princeton.