Postural Expressions of Emotion in a Motion Captured Database and in a Humanoid Robot

Andrea Kleinsmith[1], Issam Rebai[2], Nadia Berthouze[1], Jean-Claude Martin[2]

(1) UCL Interaction Centre (UCLIC)
Malet Place Engineering Building
8th floor University College London Gower Street
London WC1E 6BT
+44 (0) 20 7679 0690 (30690)
n.berthouze@ucl.ac.uk

(2) LIMSI-CNRS
BP 133
91403 Orsay, FRANCE
+33.6.84.21.62.05
MARTIN@LIMSI.FR

ABSTRACT
Humanoid robots need to be endowed with the capacity to express emotions using full body postures. In this paper we describe how a motion capture database was collected and labelled. We explain how we use it for the specification of a lexicon of affective posture to be used by a humanoid robot.

Categories and Subject Descriptors
I.2.9 [**Artificial Intelligence**]: Robotics|speech and gesture generation

General Terms
Human Factors.

Keywords
Motion capture, emotion corpora, humanoid robot, posture.

1. INTRODUCTION
Humanoid robots are developing rapidly. Few humanoid robots have been endowed with a library of postures expressing emotions. Motion capture is occasionally used to control on-line a humanoid's robot movements (http://vimeo.com/5447684). Yet, interactive humanoid robots need to generate dynamically non-verbal behaviors and thus require a lexicon of various postures expressing emotions. In this paper, we explore the usability of a motion capture database of postural expressions of emotion for the design of such a lexicon of expressive posture to be displayed by the Nao humanoid robot (http://www.aldebaran-robotics.com/).

2. MOTION CAPTURE
We used a motion capture system (VICON system) to collect 3D affective postures of 13 human subjects.

AFFINE '09, November 6, 2009 Boston, MA

The participants (called actors hereafter) were university students and hence not professional actors. 11 actors were Japanese, one was from Sri Lanka and one was from the US. The actors were instructed either in Japanese or in English. It has to be noted that the Sri-Lankan actors spoke fluently English as this was the language used in their educational system.

Dressed in a suit with 32 markers attached to various joints and body segments, each actor was asked to perform an in-place posture expressing Anger, Fear, Happiness, and Sadness. The actors were directed to perform the emotion postures in their own way, as no constraints were placed on them. We define a posture as any stance involving the hands and/or body that can convey emotions or feelings. Each affective posture was captured by eight cameras and represented by contiguous frames describing the position of the 32 markers in the 3D space. In total, 108 affective postures were captured.

In order to label the collected posture, we decided to use observers' evaluations rather than what the actor thought they were expressing as our actors were not professional. To this aim, we used the original motion capture data to build affectively expressive avatars, shown in Figure 1, by selecting the frame (i.e. a static posture) that the actor evaluated as being the most expressive instant of the posture.

We used computer avatars instead of human photos in order to create a non-gender, non-culturally specific 'humanoid' in an attempt to eliminate bias. Furthermore, by using the avatar, the subjects are not affected by facial expressions. Various studies have shown that bodily expressions do bias the perception of the affective state conveyed by the face, and vice versa [1, 2].

We used an eight word forced-choice experimental design to evaluate emotion perception through whole body posture. This method was purposely chosen as it has been widely used in studies aimed at assessing cross-cultural agreements in the expression of emotion [3].The observers consisted of 25 Japanese, 25 Sri Lankans, and 20 Caucasian Americans. The Japanese and Sri Lankan observers were all within a similar age range (University students) and came from a similar educational background, educational level, and career. The age range amongst the Americans was much greater, as was the career status, while the educational level was similar to that of the other cultures.

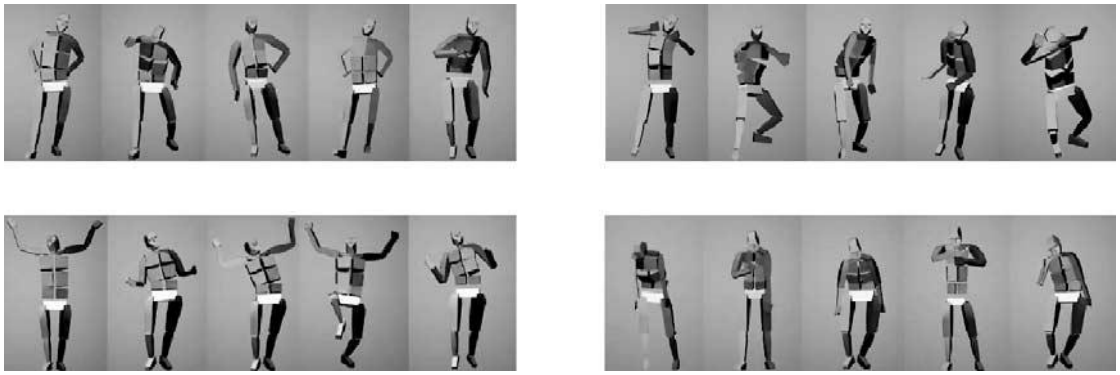

Figure 1: Examples of affectively expressive avatars for each emotion category. From left to right, top to bottom, they represent Angry, Fear, Happy and Sad body expressions.

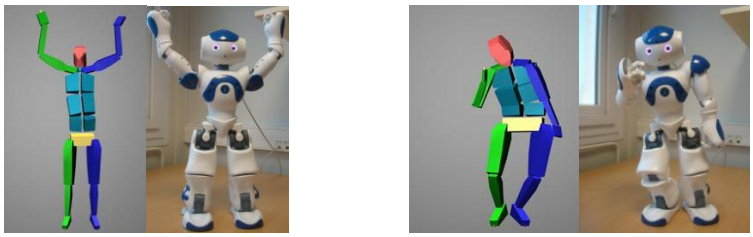

Figure 2: Examples of mapping between avatar's motion captured posture and the manully specified posture of the humanoid robot (from Left to Right: FH2_7-18_541 (best recognised as Joy by US subjects) and JS3_10-7_455 (Sad)).

The experiment was conducted online as a series of webpages. Experiment instructions were given in English for the non-Japanese observers, and in both English and Japanese for the Japanese observers. Again, as discussed above, English was used for the Sri Lankan observers because it is the main language used in their educational system. The postures were presented in a randomized order, differing for each participant. For each page (one posture per page), subjects were asked to (1) rate the intensity of the emotion, defined by a value between 1 (lowest) and 5 (highest) to indicate how emotional is the posture, and (2) choose an emotion label to represent the posture displayed based on an eight word list comprised of pairs of labels indicating two nuances of the same emotion: anger (angry, upset), fear (fearful, surprised), happiness (happy, joy), and sadness (sad, depressed). The agreement analysis between observers and actors and the cross-cultural difference analysis in evaluating the postures are reported in [4].

3. APPLICATION TO A HUMANOID ROBOT

Similarly to the avatar used in the previous section, the Nao robot does not display any facial expressions. Hence, it is of main importance to properly specify postures that are intended to express emotions. Automatic mapping between mocap data and an output device such as Nao is not straightforward (e.g. the mocap data is a set of points, whereas the postures of the Nao robot have to be specified using Euler angles).

Our approach consisted in starting from the avatar's posture which were best recognised and to try to manually specify them on the Nao robot. When not possible (e.g. because the corresponding value for an angle is not possible on Nao or would make it fall), we selected the 2nd best posture. Three postures were specified for each of the following emotion category: Anger, Sadness, Happiness, Fear. The dynamics of the postural expressions is specified using Python scripts.

4. ACKNOWLEDGMENTS

Part of this work was funded by the ANR GV-Lex project. The authors wish to thank Erwan Leroux for his work on the project.

5. REFERENCES

[1] Meeren, H. K. M., van Heijnsbergen, C. and de Gelder, B. Rapid perceptual integration of facial expression and emotional body language. *Proceedings of the National Academy of Sciences of the USA*, 102, 45 (2005), 16518-16523.

[2] Clavel, C., Plessier, J., Martin, J. C., Ach, L. and Morel, B. *Combining Facial and Postural Expressions of Emotions in a Virtual Character*. Intelligent Virtual Agents (IVA'2008). Amsterdam, The Netherlands. 16 - 18 September 2009. Springer. pp 287 - 300.

[3] Keltner, D., Ekman, P., Gonzaga, G. C. and Beer, J. *Facial expression of emotion*. Oxford University Press, City, 2003.

[4] Kleinsmith, A., De Silva, R. and Bianchi-Berthouze, N. Cross-Cultural Differences in Recognizing Affect from Body Posture. *Interacting with Computers*, 18, 6 (2006), 1371-1389.