

© 2003 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

N. Berthouze, T. Fushimi, M. Hasegawa, A. Kleinsmith, H. Takenaka and L. Berthouze, "Learning to recognize affective body postures," The 3rd International Workshop on Scientific Use of Submarine Cables and Related Technologies, 2003., Lugano, Switzerland, 2003, pp. 193-198, doi: 10.1109/CIMSA.2003.1227226.

<https://doi.org/10.1109/CIMSA.2003.1227226>

Access to this work was provided by the University of Maryland, Baltimore County (UMBC) ScholarWorks@UMBC digital repository on the Maryland Shared Open Access (MD-SOAR) platform.

Please provide feedback

Please support the ScholarWorks@UMBC repository by emailing scholarworks-group@umbc.edu and telling us what having access to this work means to you and why it's important to you. Thank you.

Learning to Recognize Affective Body Postures

Nadia Berthouze¹, Tsuyoshi Fushimi¹, Makoto Hasegawa¹, Andrea Kleinsmith¹, Hayato Takenaka¹,

¹Department of Computer Software
Aizu University, Aizu Wakamatsu, 965-8580, Japan
Phone: (+81) 242 37-2790, email: nadia@u-aizu.ac.jp

Luc Berthouze^{2,1}

² Neuroscience Research Institute
Tsukuba AIST 2, Tsukuba, 305-8568, Japan
Phone: (+81) 298 615369, email: luc.berthouze@aist.go.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. But current systems do not seem to react to users, or only in clearly pre-defined ways. In that sense, current systems miss the bi-directionality typical to human social interaction. Social interaction is characterized by a multi-channel communication, in which each actor captures and reacts to signals by the other actor. To this aim, a computer or a robot has to be able to capture and interpret signals sent by the human partner in order to achieve social interaction. One of the most important channels of communication is physical interaction. The body is used to interpret the affective state of an interlocutor. 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 learn to recognize affective states from body postures.

Keywords – Human-Machine Interaction, Affective Body Language, Incremental Learning, Categorization

I. INTRODUCTION

Robots are assuming an increasingly important role in our society. Robots are now used to entertain an active and affective communication with human agents. The dog-shaped robot, AIBO, [1] is used as an amusement toy and PARO [2] is a seal-like robot developed to help the healing process of infants in a hospital environment. These robots express emotions through various communication channels. For example, when caressed, PARO will show pleasure by raising its head, while AIBO will move its tail when it sees its owner.

During interaction, humans go beyond the messages exchanged verbally. They interpret such messages and decide on appropriate reactions by looking at other clues sent through other modalities. The same verbal message can assume different meanings, depending on contingent voice tone, facial expression or body posture. Studies show that body language is the main means of communication. Vinayagamoorthy et al. [3] claimed that nonverbal communication is important and amounts to up to 93% of communication and that 55% of it consists of body language while 38% is expressed through

tone of voice. Individual communication style, behaviors, attitudes, motivation, and personality are all culturally dependent [4]. For example, American people tend to express emotion more openly than Japanese people [5].

Many researchers have been trying to examine the relation between emotion and nonverbal cues [6][7][8][9]. In the same vein of work as [9], 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. Suzuki et al. [10] tried to maximize the interaction by embedding artificial emotions into a robot. Its emotional state dynamically changes according to a result of competition within its four internal emotional states (i.e., *happy*, *angry*, *tranquil*, and *melancholic*), and its states are exhibited through lights, music, and behaviors. Greta [11] is a conversational agent able to converse with the user by exhibiting synchronized and coherent verbal and nonverbal behavior. The agent has a personality and a social role (information delivery), and it expresses emotions through facial expressions.

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.

Various studies have been carried out to create computational models of the relation between signal and affective state. Fukuda and Kostov [12] developed a simplified human-based Emotional Model based on Wavelet/Cepstrum techniques for

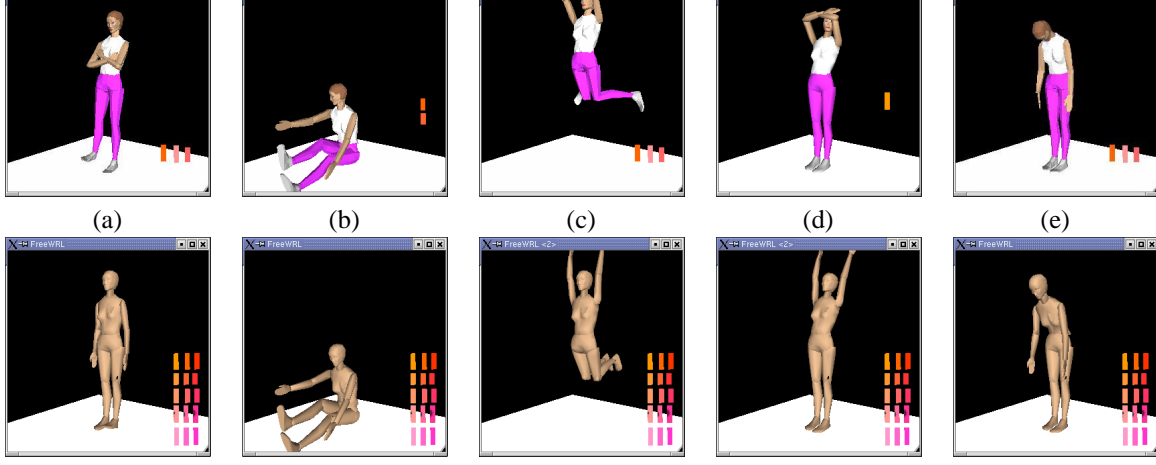


Fig. 1. Emotional animations: the first row corresponds to human-like motions, the second row corresponds to robot-like motions. (a) *irritated-angry*, (b) *tantrum*, (c) *happy*, (d) *scared*, (e) *sad*.

emotion extraction from a human voice. A system [13] was created that can recognize emotional and para-linguistic communication by combining facial expressions and voice.

Suggested by [14], 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. A car model showing such abilities has been recently launched by Toyota [15]. This car detects potentially dangerous emotional states in the driver by performing physiological measurements and by subsequently reacting appropriately to quiet him/her down.

Meanwhile, little attention has been placed on the visual signals to be extracted from body posture. Dance and choreographic studies [16] have shown that it is a powerful and frequently used means of communication. In this paper we describe our experiments to study the importance of body language in social communication. In particular, we focus on studying the features of body postures that are responsible for a given affective meaning. The results of these experiments led to the development of a computational system with the ability to incrementally learn to recognize affective messages conveyed by body postures. The modeling process is seen as a categorization process where affective messages are grounded on the description of body postures. The set of categories and body postures is not predefined but it is incrementally learnt as it is presented to the system. This approach may reflect the human learning process. In addition, it allows for associating a same affective category to various body postures. This is an important feature because body language varies between cultures, sub-cultures and even between persons within the same sub-culture [17].

The paper is organized as follows. At first, we report on our experiments. Then, we suggest an experimental scenario

and we describe in detail the architecture we have developed. Finally, we conclude with preliminary experimental results of the affective categorization process.

II. BODY POSTURE 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 [16][18] in particular, have shown that a movement conveys a different affective message when its features, e.g. its spatial dimensions, are modified. Other studies, such as described in [19], 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' ability 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 two avatars with the same human-like body but with movements having different degrees of freedom. While the first avatar could perform human-like motions, the second had restricted abilities. We modeled the second avatar by implementing the same number of degrees of freedom as in the robot used in the project. This avatar could not bend nor cross its arms, and neither could it incline its head. Facial expressions were not used so as to focus only on body motion. The two avatars could be animated by using key-framing and inverse kinematics.

lected to evaluate the posture of both avatars. We selected five emotional states: *happy*, *sad*, *angry*, *tantrum*, and *scared*. For each emotion, we prepared three variations of the same animation. The variations were based on modifications of the following features: speed, symmetry and amplitude. The same animations were reproduced with the robot-like avatar by reducing the number of degrees of freedom according to its motion abilities. In a questionnaire, the subjects were asked to:

- describe the affective state conveyed by each animation.
- identify a context or a situation that could have caused the affective state.
- evaluate the intelligibility of the affective state on a 5 degree scale (from very easy (1) to very difficult (5) to understand).
- identify the most relevant features of each animation with respect to the conveyed affective state.

Animation	Labels	C	I	S	Sp	A	R
Anim. a	angry (3)	↗	↗	↘	↗	—	—
	irritate (2)						
	impatient (2)						
Anim. b	tantrum (1)	↗	↗	↘	—	—	↗
	selfish (3)						
	spoilt child (2)						
	sacred (1)						
Anim. c	happy (2)	↗	↗	↘	—	↗	↗
	joyful (4)						
	elated (1)						
Anim. d	scared (2)	—	—	—	↗	↗	—
	hopeless (2)						
	tired (1)						
	impatient (1)						
Anim. e	sad (2)	↘	↘	—	↘	↗	↗
	hopeless (2)						
	tired (1)						

TABLE I

THE TABLE SUMMARIZES SUBJECTS' OBSERVATIONS. THE FIRST COLUMN LISTS THE LABELS USED BY THE SUBJECTS, AND THE LABEL FREQUENCY IS REPORTED IN BRACKET. THE MOTION FEATURES ARE: C = COMPLEXITY, I = IRREGULARITY, S = SYMMETRY, Sp = SPEED, A = AMPLITUDE, R = RIGIDITY. THE ARROWS INDICATE IF THE FEATURE WAS CONSIDERED, IN AVERAGE, AS NECESSARY (↗), NOT DESIRED (↘) OR IRRELEVANT (—).

A. Results and Discussion

The *irritated-angry* animation (see Figure 1-(a)) consisted of an avatar, with crossed arms and the head slightly upward, repeatedly tapping the ground with its foot. The set of *irritated-angry* animations differed only in the speed of tapping. Even if the animations could be recognized by the position assumed by the upper body, some subjects argued that the

mainly related to the rhythmic repetition of the tapping of the foot. The impossibility to manipulate the upper body of the robot-like avatar made it difficult for the animations to be recognized as such by the subjects.

The *tantrum* animation (see Figure 1-(b)) was represented by the avatar laying down on the ground with the body slightly bent upwards, and showing repetitive arm and leg movements. The *happy* animation (see Figure 1-(c)) was represented by jumps with arms moved up and down. Variations of both animations mainly involved their symmetry feature. In these animations, subjects showed a very strong rejection of symmetry and regularity of movements. Such characteristics were defined as unnatural. Even when the symmetry was removed, the regular repetition of the same movement was seen as awkward and considered more similar to a gymnastic exercise.

The *scared* animation (see Figure 1-(d)) was represented by the avatar covering its face with its arms and bending the upper body slightly backward. In the robot-like avatar, the crossing of the arms was not possible. However, in both cases, the motion was not easily recognized because of the low speed and because the body was only slightly bent backwards (limited amplitude of the motion). The rigidity of the body in the second avatar was however not perceived as unnatural.

In the *sad* animation (see Figure 1-(e)), the upper body of the avatar was made to bow slowly. In the human-like avatar, the head was bent down as well, while that was not possible in the robot-like avatar. The motion was carried out slowly in all three variations but with varied amplitudes. While the affective state was easily recognized, the inclination (or amplitude) of the bow was the key point for associating the motion to a given affective state. Animations with high degrees of inclination were discarded as resembling a gymnastic exercise rather than expressing a natural affective gesture. Another important feature was the rigidity of the robot-like avatar's arms, hands and head: this appeared to explain why the robot-like avatar was not seen as natural.

Table I summarizes the observations made by the subjects on each animation. The first column lists the labels used by the subjects when classifying the affective state conveyed by each animation. The number in brackets indicates the number of subjects that used such label. From the user's observations, we identified as relevant the following motion features: Complexity (C), Irregularity (I), Symmetry (S), Speed (Sp), Amplitude (A), Rigidity (R). The arrows indicate if the features were, on average, considered by the subjects as necessary (↗), irrelevant (—) or not desired (↘). 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 the *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

of the gesture and for effectively conveying an affective state. Rigidity not only resulted in the motion being effectively seen as unnatural but also decreased the intelligibility of the affective state.

We computed the average intelligibility values associated by each subject to each animation of the first avatar. The average values were: *angry* 1.43, *tantrum* 2.69, *happy* 1.86, *scared* 2.57 and *sad* 2.00. The *angry*, *happy*, and *sad* animations showed to be, on average, easy to recognize, while *scared* and *tantrum* animations showed to be more ambiguous. Instead, motions of the robot-like avatar were generally considered as awkward. One explanation of this phenomenon is that of expectation. Because the robot-like avatar had a human-like body, subjects were expecting human-like motions. We believe that motions would have been better accepted if the avatar body had been more robot-like.

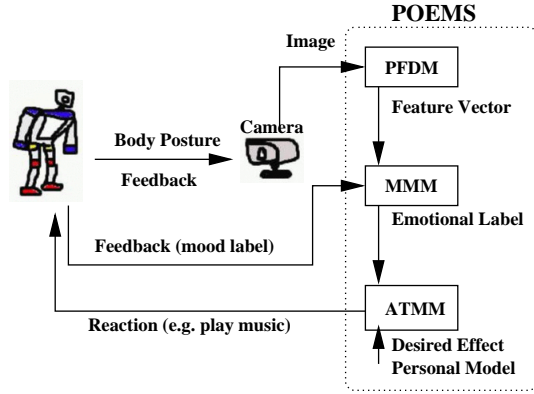


Fig. 2. The architecture of the POSture-Emotion Mapping (POEM) System.

III. SCENARIO

A video camera captures a human made expressive posture and sends it to the computational system. The system analyses and maps the posture 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. Figure 2 depicts the architecture of our POSture-Emotion Mapping (POEM) system. The architecture consists of 3 main modules: the Posture Feature Detection module (PFD), the Mood Mapping module (MM), and the Affective Transition Modeling module (ATM).

To facilitate the study of the relation between posture and emotion, we used a small humanoid robot (see Figure 3-a) whose joints have been color marked. The robot's expressive posture is captured by the video camera, and the image is sent to the PFD module. The PFDM, using image processing algorithms, outputs a vector of values that describes the captured posture. The posture description process is depicted in Figure 3. To overcome problems associated with zooming, the values are normalized using the distance between the centroids of

description vector of 16 features:

- joint angle of the right/left knee.
- joint angle of the right/left hip.
- joint angle of the right/left shoulder.
- length of the right/left upper leg.
- length of the right/left lower leg.
- length of the right/left arm.
- length of the right/left shoulder.
- length of the torso.
- length between hips and the ground.

The MM module maps the posture description vector into an emotional label. The emotional label is used by the ATM module to select the action to be performed. Our current ATM module selects and plays pieces of music to trigger a mood transition in the partner. The choice of the piece of music is based on a desired transition (which we can call motivation or goal) and a personal mood-transition model. This model will be described elsewhere. To improve its performance, the POEM 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 MM module to explicitly indicate the correct emotional label when the MM outputs the wrong label. The second type of feedback is postural feedback and corresponds to changes in the user's body posture as a reaction to the musical piece. The POEM system can hence check if the mood transition was the one predicted by the transition model endowed into the ATM module.

In the next section, we focus on the description of the MM module for the incremental learning of affective postures.

IV. CATEGORIZATION PROCESS

The mapping of posture features into emotional labels can be seen as a categorization problem. We propose to use an association network, called Categorizing and Learning module (CALM) [20], 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. When most artificial networks suffer from a variety of problems like a lack of speed, lack of stability, inability to learn either with or without supervision, inability to both distinguish between new and old information and generalize over patterns, CALM has improved upon these deficiencies by incorporating structural and functional constraints related to the modular structure of both the human brain and its information processing, such as modularity and organization with excitatory and inhibitory connections.

The idea of modular architecture leads to an increased stability of representations and a reduced interference by subsequent learning because of the reduced plasticity. CALM has an attentional control mechanism that is sensitive to the novelty of the input pattern. It also has two learning modes, namely elaboration learning mode and activation learning mode. Elaboration learning is used to learn new information and quickly create new categories. Activation learning is used to preserve the memory of old information by strengthening existing associations.



Fig. 3. Posture description process: (a) original image captured by the camera, (b) resulting image after color segmentation, (c) removal of noise labelling, (d) 11 joints labeled according to their type, (e) 11 joints colored according to the side they belong to, (f) resulting body skeleton.

As shown in Figure 4, a CALM module consists of four types of nodes: representation nodes (R-nodes), veto-nodes (V-nodes), arousal node (A-node), and external node (E-node). The number of R-nodes and V-nodes equals the number of inputs to the module. Information is sent to the R-nodes from other modules and the R-nodes activate the V-nodes and the A-node. The V-nodes receive activations from all R-nodes and inhibit other V-nodes, all R-nodes, and the A-node to trigger competition inside a module.

The amount of activation of the A-node indicates the novelty of the input pattern and determines the level of competition. A new input pattern triggers high competition to create a more stable representation, but a familiar input pattern triggers a decrease in competition. The A-node also determines the activation of the E-node. The E-node sends random activation pulses to all R-nodes to select one winning R-node. In the event of a deadlock, the E-node also controls the learning rate in a module as a function of the activation of the A-node. The E-node changes to a high learning rate for a new input pattern, and to a low learning rate for old input patterns. By adjusting the learning rate, the inter-weights are modified so that the CALM architecture ensures the long term storage of the pattern.

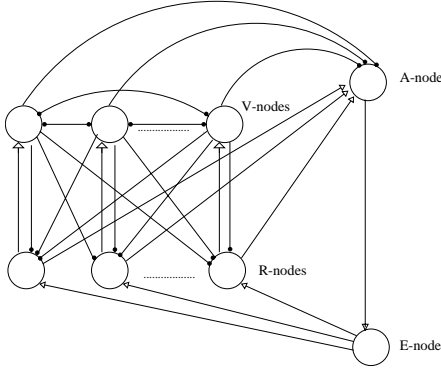


Fig. 4. A CALM module consists of four types of nodes: representation nodes (R-nodes), veto-nodes (V-nodes), arousal node (A-node), and external node (E-node). The number of R-nodes and V-nodes equals the number of inputs.

V. MOOD MAPPING MODULE: CAPTURING EMOTIONAL STATES

The Mood Mapping module consists of a CALM network whose topology is described in Figure 5. It has 1 input module,

3 internal modules, 1 output module, and 1 feedback module. The input module has 16 R-nodes. They are fed with the posture description vector. The internal modules are of different sizes. The output module corresponds to the set of emotion categories that can be recognized. The output module initially has only 2 R-nodes but it is incrementally grown as new categories are introduced. We use a modified version of CALM proposed by [21]. It includes a feedback module to supervise the learning process. During the learning phase, the feedback module forces the categorization when the system produces a wrong output.

The learning process uses a short term memory of 10 patterns corresponding to the last 10 posture descriptions learnt. Every time the MM module learns a new posture-emotion relation, it also rehearses the patterns in the short term memory. The new pattern is then added to the short term memory and the oldest pattern is removed. In other words, the mapping system incrementally learns new pattern-emotion relations and at the same time reinforces its closest past experiences without completely forgetting the oldest ones.

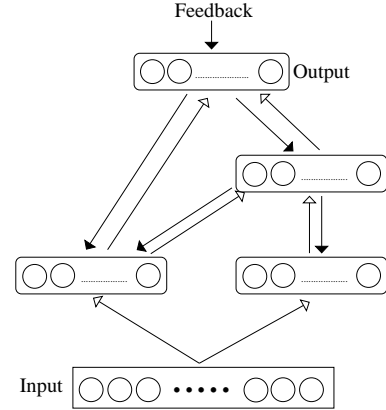


Fig. 5. The architecture of the CALM network with feedback used in the Mood Mapping module. The input module receives 16 input values, which describe the posture to be labeled. The size of the output module increases incrementally when new affective categories are introduced.

A. Experimental Results

We tested the Mood Mapping module with 108 images involving affective postures. The set included 36 different postures for each type of affective state, i.e. *happy*, *angry* and *sad* postures. The 108 images were split into two sets: 54 were

phase. 54 images were randomly fed to the Mood Mapping module.

Figure 6 shows the unfolding of the learning process and the emergence of categories over time. The horizontal axis denotes the number of learning steps, and the vertical axis represents the cumulated recognition percentage of each category. After a few learning steps, only one category (*sad*) emerges, and every posture is mapped onto that category. After a few more steps, the system started to discriminate two simultaneous categories, *sad* and *happy*. This is highlighted by the decrease of the recognition curve of the *sad* category and the increase of the recognition curve of the *happy* category. Soon after, the module recognized the presence of the third category (*angry*). As a result, the recognition curve of *sad* shows an oscillating trend (indicating competition), while the *angry* curve started to increase (successful discrimination). The learning process for the 54 images required about 1000 steps to reach the stop condition, i.e. all patterns were recognized correctly upon 2 consecutive presentations.

We tested the trained CALM network on the 54 posture descriptions of the testing set. Only 1 error occurred: the posture was related to a *sad* emotion but was classified as *angry* by the model.

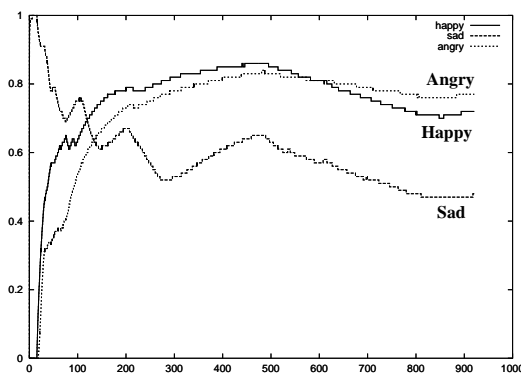


Fig. 6. Unfolding of the learning process of three affective categories. The horizontal axis denotes the number of learning steps. The vertical axis denotes the cumulated recognition percentage.

VI. CONCLUSIONS

One original aspect of this project resides in its treatment of communication. Namely, we see human-machine communication as a multi-modal language where words (expressed in any modality) contain instances of affective categories that emerge from competitive signals describing the body posture of a human subject. A system can be considered to be capable of affective communication when (a) it has the ability to associate a certain modal message to an affective message or state, and (b) it has the ability to convey, through one or many modalities, an affective state or message. In this respect, one expected result of our study is a clarification of the characteristics of the body language modality as a means for emotional communication, and how the emergence of affective categories could be

an active actor in the interaction with the human. The ability to recognize emotions is a requirement for social interaction to take place. Up to now, body language has not really been explored, with the focus put on voice and facial expressions. To increase the complexity of the possible postures to be modeled, we are now exploiting motion features by analyzing data acquired with a 3D motion capture system using 38 sensors. By shifting to human generated motions (rather than robot), we are ready to consider the time dimensions and the related motion features already detected as relevant in our preliminary experiments.

References

- [1] AIBO, "Entertainment Robot", <http://www.aibo.com/>, 2003.
- [2] T. Shibata, T. Tashima, and K. Tanie, "Emergence of Emotional Behavior through Physical Interaction between Human and Pet Robot", *Proceedings of the International Workshop on Humanoid and Human Friendly Robotics*, VI-4, pp. 1-6, 1998.
- [3] V. Vinayagamoorthy, M. Slater, and A. Steed, *Emotional personification of humanoids in Immersive Virtual Environments*, tech. report Equator-02-029, Dept. Computer Science, Univ. College London, Sept. 2002.
- [4] P.O. Brown, "Setting the Stage for the Culturally Adaptive Agent", *Proc. AAAI Fall Symposium on Socially Intelligent Agents*, pp. 93-97, 1997.
- [5] D. Matsumoto, T. Consolacian, H. Yamada, R. Suzuki, B. Franklin, S. Paul, R. Ray, and H. Uchida, "American-Japanese cultural differences in judgments of emotional expressions of different intensities", *Cognition and Emotion*, N. 6, pp. 721-747, 2002.
- [6] N. Bianchi-Berthouze, C. Lisetti, "Modeling multimodal expression of users's affective subjective experience", *International Journal on User Modeling and User-Adapted Interaction*, V. 12, N.1, pp.49-84, 2002.
- [7] B. Kort and R. Reilly, "Theories for Deep Change in Affect-sensitive Cognitive Machines: A Constructivist Model", *Educational Technology & Society*, vol. 5, no. 4, pp. 56-63, 2002.
- [8] K.J. Overbeeke, P. Vink, and F.K. Cheung, "The emotion-aware office chair", *Proc. Int'l. Conf. Affective Human Factors Design*, Asian Academic Press, London, pp. 262-267, 2001.
- [9] C. Breazeal, B. Scassellati, "How to build robots that make friends and influence people", *International Conference on Intelligent Robots and Systems*, Kyongju, Korea, 1999.
- [10] K. Suzuki, A. Camurri, P. Ferrentino, and S. Hashimoto, "Intelligent Agent System for Human-Robot Interaction through Artificial Emotion", *Proceedings of IEEE Int'l. Conference on Systems Man and Cybernetics*, San Diego, 1998, pp. 1055-1060.
- [11] C. Pelachaud, V. Carofiglio, and I. Poggi, "Embodied Contextual Agent in Information Delivering Application", *Int'l. Joint Conference on Autonomous Agents & Multi-Agent Systems*, Italy, pp. 758-765, 2002.
- [12] S. Fukuda and V. Kostov, "Extracting Emotion from Voice", *Proceedings of IEEE Int'l. Conference on Systems Man and Cybernetics*, Japan, vol. 4, pp. 299-304, 1999.
- [13] Y. Tian, T. Kanade, and J. Cohn, "Recognizing action units for facial expression analysis", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 23, No. 2, February, pp. 97-115, 2001.
- [14] R. W. Picard, "Affective Computing", MIT PRESS, 1997.
- [15] "Car with feelings launched by Toyota", http://www.ananova.com/news/story/sm_427116.html, 2003.
- [16] R. von Laban, "The mastery of movement", Princeton, 1988.
- [17] "Embodied Conversational Agents", eds J. Cassell, MIT press, 2000.
- [18] A. Camurri, S. Hashimoto, K. Suzuki, and R. Trocca, "Kansei analysis of dance performance", *1999 IEEE International Conference on Systems, Man and Cybernetics*, Japan, pp. 327-332, 1999.
- [19] T. Nakata, "Generation of Whole-Body Expressive Movement Based on Somatological Theories", *Proc. 2nd Int'l workshop on Epigenetic Robotics*, pp. 105-114, 2002.
- [20] J.M.J. Murre, R.H. Phaf, and G. Wolters, "CALM: Categorizing and Learning Module", *Neural Networks*, vol. 5, pp. 55-82, 1992.
- [21] L. Berthouze and A. Tijsseling, "Acquiring ontological categories through interaction", *Proc. Fifth Int'l Conf. Human and Computer (HC-2002)*, pp. 160-166, 2002.