



HAL
open science

FIRST EVALUATION OF EMI MODEL OF INTERACTION

Sébastien Saint-Aimé, Brigitte Le Pévédic, Dominique Duhaut

► **To cite this version:**

Sébastien Saint-Aimé, Brigitte Le Pévédic, Dominique Duhaut. FIRST EVALUATION OF EMI MODEL OF INTERACTION. Proceedings of the 14th IASTED International Conference on Robotics and Applications – RA 2009, Nov 2009, Cambridge, Massachusetts, United States. pp.263. hal-00507631

HAL Id: hal-00507631

<https://hal.science/hal-00507631v1>

Submitted on 30 Jul 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

FIRST EVALUATION OF EMI MODEL OF INTERACTION

ABSTRACT

This article presents research work done in the domain of nonverbal emotional interaction for the EmotiRob project. It is a component of the MAPH project, the objective of which is to give comfort to vulnerable children and/or those undergoing long-term hospitalisation through the help of an emotional robot companion. The studies carried out on perception and emotional synthesis have allowed us to determine a method of emotional expression for our public. It is important to note that we are not trying to reproduce human emotion, but trying to make a robot emotionally expressive. Our experiments on the subject have helped us determine the minimum number of degrees of liberty necessary for the robot and helped us answer some question to validate our emotional model in a next experimentation. We begin the article by MAPH and EmotiRob project presentation. Then, we quickly describe the computational model of emotional experience *iGrace* that we have created, as well as give its evaluation. We conclude with a description of the architecture of Emi, as well as improvements to be made to its third generation.

KEY WORDS

Emotion, interaction model, companion robot, simulation.

1. Introduction

Since a few years, the new challenge is to build systems that will offer behavior enrichment through their interaction with humans. Facial expressions play a major role in the coordination of human conversation [1] and constitute an essential method of human communication.

Robotherapy tries to apply the principles of social robotics with the added goal of improving the psychological and physiological state of those who are sick, secluded, or suffer from physical or mental handicaps. The robots seem to be able to play a part in accompanying and awakening them, and for that purpose should be equipped with a maximum of capacities of communications.

The experiments in this field of robotics – with robots like Paro [2,3,4], i-Cat [5], Cosmobot [6], Kismet [7], Kobie [8], etc. – clearly showed that robot companions can give a certain amount of moral and psychological comfort to

fragile people.

2. EmotiRob project

In this context we began an experiment [9] using the Paro robots to check whether or not the reaction / interaction with the robots was dependent on cultural context. These experiments showed us two principal directions in which it would be interesting to work. The first one deal with mechanical problems: the robot must be very light, easy to take and to handle, easier than Paro; moreover, it must have a great deal of autonomy. The second one leads to changing the man-machine interaction: the psychological comfort that the robot can provide is related to the quality of the emotional tie that the child has with it.

The MAPH project [10] objective is to give comfort to vulnerable children, from 4 to 6 years, and/or those undergoing long-term hospitalization with the help of a robot which can be used as an emotional companion in daily life. EmotiRob which is a sub-project from MAPH aims to design this robot to make an emotional interaction between the children and his companion. We want to equip this robot with the necessary capacities of perception and comprehension of natural language so that it can build a formal representation of the emotional state of its interlocutor. The project also comprises the design of a model of the emotional states of the robot and their evolution, in order to make its reactions seem as natural as possible.

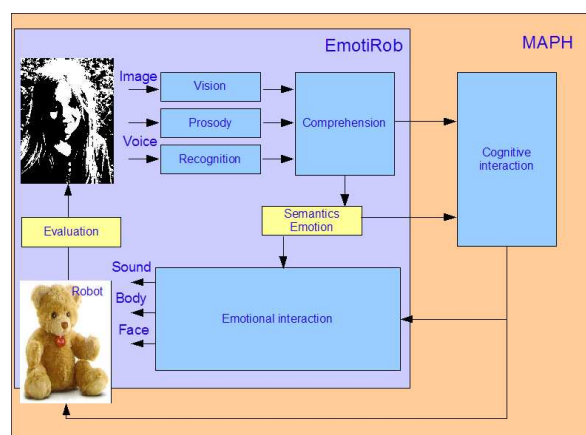


Fig. 1. Project synopsis

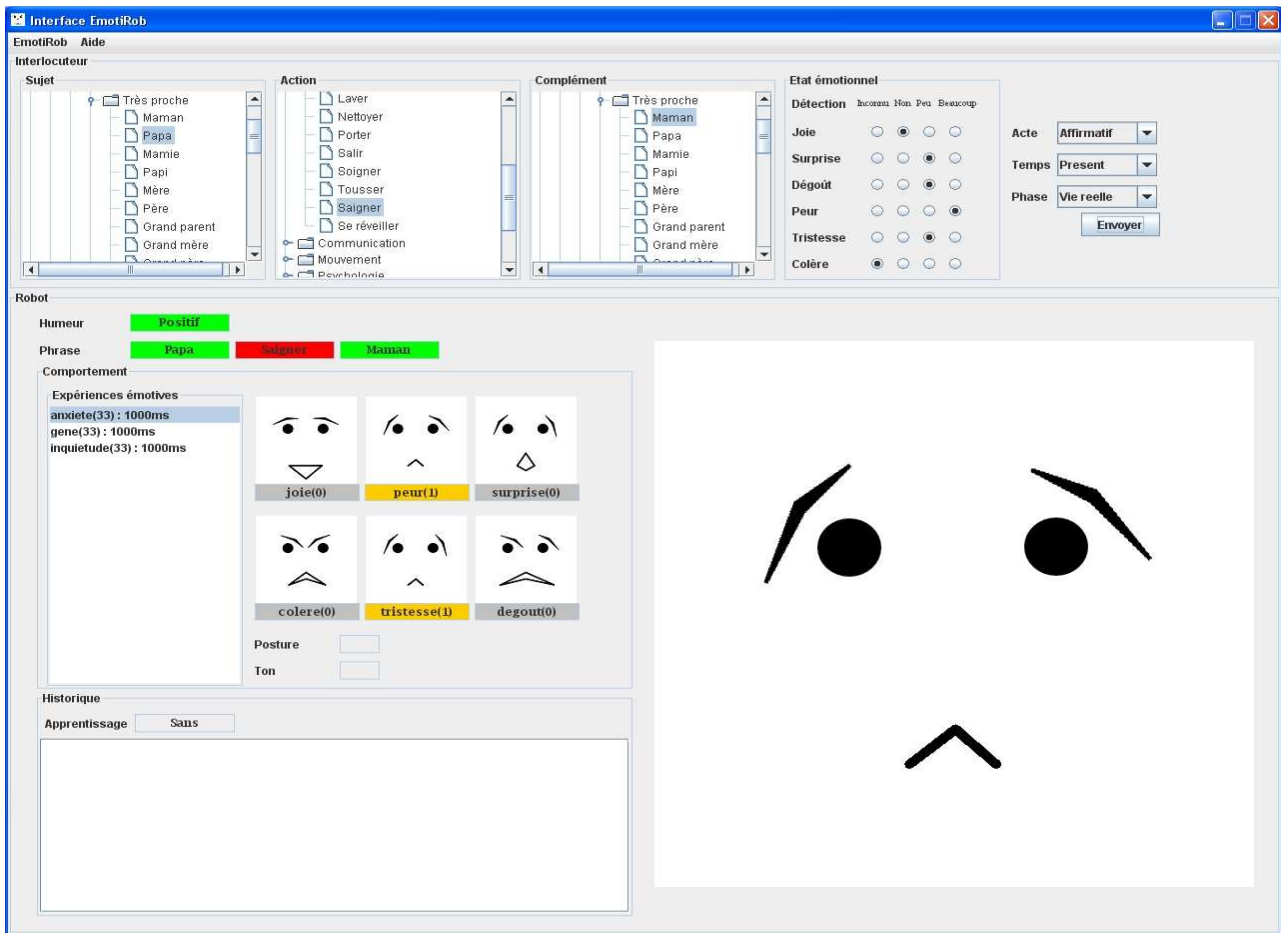


Figure 2. Simulation interface

The early stages of the project first allowed us to define the necessary degrees of freedom [9] for the robot to express primary emotions, then to define an emotional model [11] for non-verbal and emotional interaction with the interlocutor. We had to take into account the difficulties arising from the understanding of speech, do various manual annotations between emotions, emotional experiences, the words, speech and personality of the robot. The instantiation of the model ended, it is now possible to start the experiments needed to validate it and to check if the acceptance rate of the interaction was acceptable. The great question we were asked was whether the behavior linked to the speech was coherent, if the emotions expressed were recognized, and whether the dynamics of emotions seemed natural. In addition, we had to experience the speed of the expressions and the behavior length in order to include them in the robot. We will try to answer all these questions in our evaluation

3. Simulation Interface

The robot having reached the phase of integration and testing, we had to conduct a test interface (see Fig. 2) to begin the experiments. Not only can this interface enter the necessary information input, but it can also simulate the emotional state and behavior required of the robot according to the information it provides. This interface communicates with another application that is the heart of the system, or rather the application that will be mounted on the robot (see Internal Architecture Simulation). It must be stressed that the test interface allows only one simulation input / output and not data processing.

The behavior simulation is a display of the characteristics selected for interaction for each emotional experience. Each experiment is characterized by emotional facial animation, the description of the posture and tone of voice, and a length of simulation. The emotional experience can be a mixture [12] of different emotions; it is not unlikely that there will be a mix of facial expressions to animate the face. Emotions representing

emotional experiences are characterized by a coefficient and a color code:

- Gold: emotion displayed
- Silver : emotion not displayed

The interface also shows the robot's mood and the various affects associated with words of speech through two colors:

- White: neutral
- Green: positive.
- Red: negative.

Finally, in learning mode, the robot displays a history of behavior that it may have for the same speech. This information is not especially designed for the user, but is very useful to us because it allows us verify if the learning takes place properly, and also to check that the behavior of the robot is not always the same. A phenomenon of repetition over time will quickly frustrate the caller, especially if it is a child.

4. Internal Architecture Simulation

To realize emotional synthesis, the first step is to establish the necessary information to understand the environment (including the interlocutor). It is important for a robot to know how to physically express itself if it is to have non-verbal communication. However, because the interlocutor is able to communicate verbally, it is necessary to understand the main words in his/her discourse, his/her intonation, as well as his/her facial expression to grasp the emotion unveiled. We would like to be able to gather the following information:

- Discourse: even though the current systems cannot perfectly understand the range of human vocabulary and especially the discourse of an interlocutor, if the context is taken into account, some words can be processed allowing for emotional reaction. The comprehension module will enable the processing of these data to be analysed, allowing us to gather a series of information, such as the subject that is doing the action, the action, the object or subject that is undergoing the action. Moreover, by combining the acts of language, times of action, and emotional state of the interlocutor, it is possible to react without completely understanding the discourse. This reaction can still be coherent with what is said.
- The sound signal: similar to the video signal, there are 2 uses for sound. The first is to reinforce the decision taken about the emotional state of the child while speaking, as prosody is not the same for the different emotions felt while speaking. The second case is for system protection. Depending on the level of sound intensity, an appropriate reaction may be taken

by the system.

- The video signal: this information is useful for 3 cases. The first is to be able to follow the child's face. During a conversation, the interlocutor could quickly be disorientated or may think that the conversation lacks interest if his/her partner is not looking at him/her. The second case allows for the affirmation of the emotional state of the child thanks to recognized facial expression that is associated to an emotion. Finally, the signal will be very helpful in emergency situations. As children sometimes shake their toys rather roughly, it would be necessary to stop the robot from functioning during wrongful manipulation. Thus, depending on the obstruction of the camera's field of vision, the system can react in the appropriate way and will automatically go in standby mode if necessary to avoid any mechanical mishaps.

The emotional model used is entirely based on GRACE [13] and instantiates the different modules for the project. *I-Grace* (see Fig.3) is divided into 3 distinct communicating parts: *inputs*, *emotional interaction* and *outputs*. We will describe in this document the robot's information processing and behavior simulation.

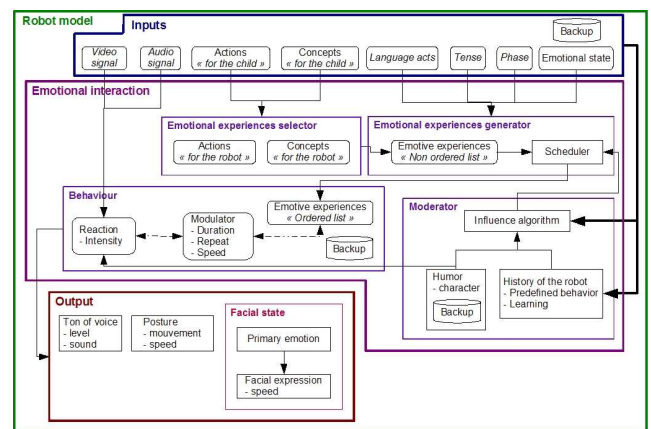


Fig. 3. Emotional interaction synopsis

4.1 Inputs

The inputs we use for our experiments are first fully simulated, but we find many of the necessary elements to make a good interaction. The users have to give, for each predefined sentence, their emotional state for the speech. They will have the possibility of transmitting incomplete or inconsistent data.

We have deliberately excluded from the module input the sound and video signal that the robot will receive. Indeed, this experiment is primarily designed to focus on the interaction between the robot and its user, based on

speech and its emotional state. Moreover, algorithms for image processing and sound are still being developed. They will be integrated into the project by 2 months.

4.2 Emotional interaction

The interaction module is like the heart or brain of the system. It receives input information, processes it with different algorithms, and determines the robot behavior that must be based on the speech. The treatments are located in the following sub-modules:

- 1) **The selector:** will weigh an initial list of emotional experiences [14] linked to the words of a speech. Each word, action or concept is associated with a vector of primary emotions. We can derive the emotional experiences associated with these emotions. We decided not to take into account emotions with a coefficient greater than 0. This factor will give a weighting of the list of experiments founded.
- 2) **The moderator:** will weigh assess a second list of emotional experiences according to robot's personality and mood. Using definitions of psychological type and MBTI model [15] for the robot, we associated some emotional experiences in the different dimensions of MBTI. The weighting of this list is by increasing the ratio of emotional experiences having an affect (positive or negative) similar to the robot's mood.
- 3) **The emotional experiences generator:** this module performs the most important interaction treatment. According to two lists of emotional experiences which we will merge, the mood and personality of the robot, the emotional state of the child, and the context for the speech (playing, imagination, real life, etc.), the algorithm will recalculate the weighting of the various lists. There are still some key points to consider.

The situation or context of the action is the first element of the list. A child at play tends to be incoherent, and sometimes even violent. We must therefore give it more freedom, and not rely on everything said that is said. On the contrary, in order to participate in his game, it is better to have a child-like emotional state while retaining our personality. We should also be careful not to be overwhelmed by overly violent words.

The second important concept is the action given by the speaker. The importance of the speech is that it reflects the action performed and the object or person to whom the action applies. That's where we see the need for a context (first item on the list).

We will end with the emotional state. The aim of this project is to comfort the child and to make him understand that his companion understands what he feels and shares his emotions. But we must also show

that it has something to say.

Other parts of speech are not useless, on the contrary, they will provide a weighting of the list, more or less aligned with the choice made from the 3 elements listed above.

- 4) **Behavior:** once the list of emotional experiences weighted, we will have to choose how this behavior is to be performed. For the moment, we only have information to simulate the experiences, and, implicitly, the emotions related to them. The first thing is to select the best experiences in the list. To avoid repetition, the first thing to be done was to filter the emotional experiences that had already been used for the same discourse. A historical base of behaviors associated to the discourse would help in this process.

For this experiment, we limited the number to 3. This will avoid excessively prolonged behavior patterns, and especially provide dynamic simulation. When the weighting is the same, the history allows a filtering process through this list to exclude the experiences that have been used recently, or for this speech. This filtering allows us to avoid behavior repetition.

To increase the interaction, and allow the robot to transmit its emotion better, we need to add for each experience a posture and a sound linked to a tone. Regarding the sound emitted by the robot, we chose to use the combination of piano notes to emotions [16]. This scientifically validated method was simple to use and very light in terms of data storage. For posture, we associated every emotion to a range of angles for positioning. The weighting of the emotional experiences list will allow a play on the angle of the torso.

4.3 Outputs

First, for a use with the robot's motors, it can be summed up by using a graphic panel that displays the emotions of the robot. In the near future, this module will manage all the robot's motors and actuators in order to synchronize the interaction.

From the final list of emotional experiences, *Behavior* sub-module gives us posture and sound to play for each item of this list. Posture is characterized by a body movement and facial expressions. The different positions that motors have to take for each posture and facial expressions are stored in a data base of patterns we manually predefined.

The motors used for movements are AX-12 Dynamixel [20] using an FTDI API [21] for communication with PC. This motor allows us play some sounds in one tone too.

5. Experimentation

The final project is geared to children with physical or

psychological handicaps. The major difficulty with this target is not being able to perform these tests as often as we wish. We will arrive in different hospitals, in partnership with the EmotiRob project, with our companion robot in its nearly-completed version. To solve this problem, we chose to do 3 sets of tests on different audiences:

- 1) The first experiment was carried out with a very large audience of all ages to gather as much information as possible on how to improve our interaction. After analyzing the results, some improvements were made. For this experiment, only the interface simulation was used.
- 2) The second experiment will begin is actually conducted with schoolchildren from 4 to 6 years of age and not suffering from disabilities. For these tests, the robot will be used, but entries will still be simulated. We will use the Wizard of Oz experiment to capture data before the final integration of the voice recognition system, scheduled for late august.
- 3) The third experiment will be conducted in Kerpape rehabilitation center with children with disabilities, and in an educational institute. These tests will be decisive for us and will make scale tests. The test conditions will be optimal to have a maximum rate of recognition for our voice and video system.

Without the validation of the first experimentation, we can't start the next.

Another positive aspect of performing experiments with a different audience each time is to analyze the subsequent perception of emotional interaction between childhood and adulthood. The next stage for the EmotiRob project will be its use with elderly.

5.1 Protocol

The first step was performed with a wide audience; it was not very difficult to find volunteers. We had made the same experimentation with 20 people working in the project, but this study can't be taking into account because they are not objectives and the results can be hedged. Finally we limited the sessions to 10 people because as explained above, this public is not the final target and the objective wasn't to modify the interaction based on comments made by an adult audience. The users were PhD French students in computer sciences old from 23 to 28 years.

The first thing the participants were asked was to disregard the fact that the interface was the face and behavior of the robot, and that all the rest (input mode, ergonomics, etc.) was not evaluated. In addition, it was necessary for these people put themselves more or less in the place of the interlocutor in order to obtain valuable feedback.

We had previously defined a list of 4 phrases that should be the basic testers. For each, we associated the following language information:

- Action tense: present.
- Language act: affirmative.
- Context of speech: real life.

This system enabled us to save the valuable time each person would take up to make his choice. The sentences given were:

- Mummy, to cuddle, daddy.
- Tiger, to attack, grandma.
- Baby, to cry.
- I, to tickle, Sister.

5.2 Evaluation grid

After distribution and explanation of evaluation grids, each person was first asked to perform the following steps:

- 1) Give the affect (positive, negative or neutral) for each word of the speech.
- 2) Give their emotional state in reaction to the speech.
- 3) Predict the robot's emotional state in answer to this speech.

This step, although easy to implement took some time to grasp because some people had trouble expressing their feelings. After entering this information, we could start the simulation for each of the sentences. We asked users to pay attention to the expression of the robot because they could not be repeated. After observing the behavior of the robot, users had to complete the following information:

- 1) What were the recognizable emotions in the robot's behavior; how would they assess its intensity: certainly not, a little, a lot, I don't know.
- 2) The average speed of expression and the behavior length.
- 3) Did you see a combination of emotions? Yes or no.
- 4) Did the emotional sequence seem natural? Yes or no.
- 5) Are you satisfied with the robot's behavior? Certainly not, a little, a lot.

5.3 Results

The objective of this experiment was to evaluate the recognition of emotions through the simulator, and especially to determine if the response the robot will give to the speech was satisfying or not.

As regards the rate of appreciation of the behavior for each speech (see Fig. 4), we observed that all the users found the simulator's response coherent, and thereafter admitted that they would be fully satisfied if the robot was

as they were expected. The fact that testers answered about the expected emotions had an influence on overall satisfaction.

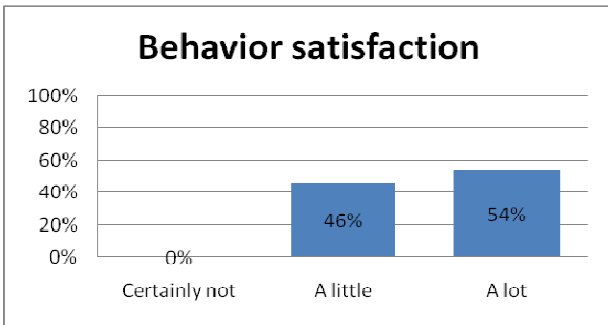


Fig. 4. Behavior satisfaction.

For the rate of emotions recognition (see Fig. 5), the figures were very satisfactory and allowed us to prepare the next evaluation on the classification of facial expressions for each primary emotion. Not all emotions are on the graph because they bore no relation to the sentences chosen. We have also been able to see that even if the results were still rather high, there were some emotions which were recognized although they were not expressed. This confirms the need to classify, and especially the fact that each expression can be a combination of emotions. The next question is to know if the satisfaction rate will be the same with the robot after the integration of the emotional model.

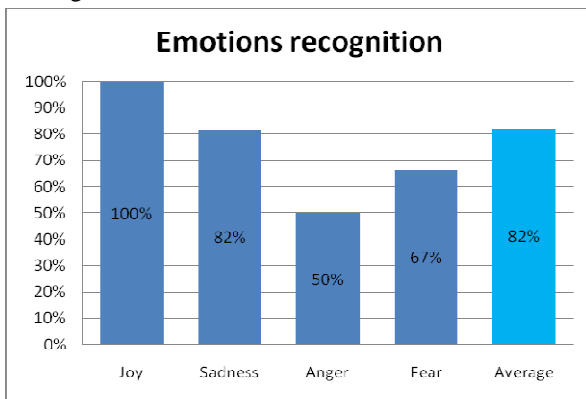


Fig. 5. Emotion recognition.

The other results (see Fig. 6 and 7) are not described in details but are used for the model integration on the robot:

- Speed of expressions: normal with 63%
- Behavior length: normal with 63%

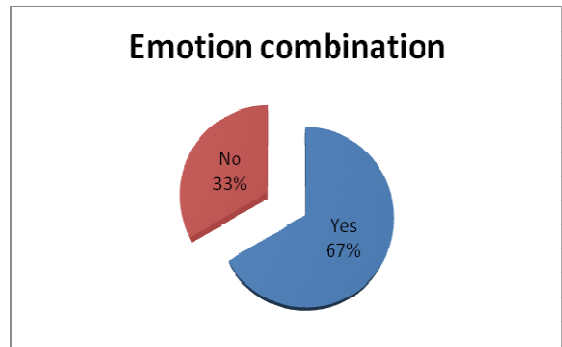


Fig. 6. Emotion recognition.

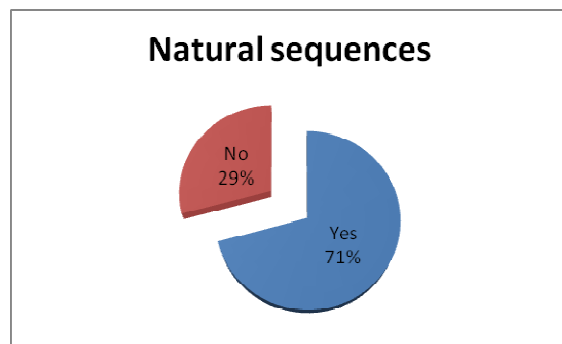


Fig. 7. Emotion recognition.

6. Robotics Conception

EmI (see Fig. 8, 9 and 10) is still in the construction and integration phase. The robot's mechanics is manufactured by the CRIFF [17], but not its implementation. We are not currently presenting aspects robotics, but giving a brief description of the work already done. Currently, we have it completely covered with plush but not fully padded. It is equipped with a CMUCAM 3 camera and a fit pc slim [18] (see Fig 11).

In a short time, all experimentation carried out with the simulation interface could be achieved with a real companion robot.



Fig. 8. First version of EmI.



Fig. 9. EmI surprise.



Fig. 10. EmI joy.



Fig. 11. Fit pc slim.

7. Conclusion and perspectives

The first experiment allowed us to answer some partially unanswered questions, particularly on the speed of expressions and behavior length. But more importantly, it confirmed the choices we had made for information processing and the behavior that we associate with the speech of an interlocutor. To fully validate the model, we have to start a new large-scale experimentation with spontaneous sentences and the simulator. The number of 10 users with predefined sentences is not significant enough.

Furthermore, we have begun in-depth research on the dynamics of emotions, in order to increase the fluidity of movement and make the interaction more natural.

The next experiment is with the robot, and will enable us to know if his expression recognition will be as high as that of the simulator after integration.

Acknowledgements

EmotiRob is currently financed by the regional council of Martinique, for the development of the emotional synthesis, the regional council of Bretagne for language comprehension, and the ANR for the construction of the robot.

Most of all, we would like to thank the regional council of Martinique as well as the ANR for their collaboration and the financing of future work.

The authors wish to thank all participants for their time spent on this experiment and their constructive comments.

References

- Cybernetics, SMC 2007. Montréal, Canada, Octobre 2007.
- [1] E. Boyle, A. H. Anderson, and A. Newlands. "The Effects of Visibility on Dialogue in a Cooperative Problem Solving Task", *Language and speech*, vol. 37, 1994, p.1-20.
- [2] K. Wada, T. Shibata, T. Saito, K. Tanie. "Effects of Robot-Assisted activity for elderly people and nurse at day service center". In *Proceedings of the IEEE*, Vol. 92, No. 11, 2004, p. 1780-1788.
- [3] T. Shibata. "An overview of Human Interactive Robots for psychological enrichment". In *Proceedings of the IEEE*, Vol. 92, No. 11, 2004, p. 1749-1758.
- [4] <http://www.parorobots.com/>
- [5] <http://www.hitech-projects.com/icat/>
- [6] C. Lathan, A. Brisben, and C. Safos. CosmoBot levels the playing field for disabled children. *Interactions -- Special Issue: Robots!* 12, 14-16, April 2005.
- [7] <http://www.ai.mit.edu/projects/humanoid-robotics-group/kismet/kismet.html>
- [8] C. Park, J. Ryu, S. Kang, J. Kim, J. Sohn, and H. Cho. *The emotion expression robot through the affective interaction: KOBIE*. In proceedings of the 1st International Conference on Robot Communication and Coordination - ROBOCOMM. Athens, Greece, 2007.
- [9] S. Saint-Aimé, B. Le-Pévédic, D. Duhaut. Building emotions with 6 degrees of freedom. In 2007 IEEE International Conference on System, Man, and Cybernetics, SMC 2007. Montréal, Canada, Octobre 2007.
- [10] S. Saint-Aimé, B. Le-Pévédic, D. Duhaut, T. Shibata. Emotirob: Companion Robot Project. In IEEE RO-MAN 2007, 16th International Symposium on Robot and Human Interactive Communication. Jeju Island, Corée du Sud, Août 2007.
- [11] S. Saint-Aimé, B. Le-Pévédic, D. Duhaut. EmotiRob: an emotional interaction model. In IEEE RO-MAN 2008, 17th International Symposium on Robot and Human Interactive Communication. Munich, Allemagne, Aout 2008.
- [12] M. Ochs, R. Niewiadomski, C. Pelachaud, D. Sadek. *Expressions Intelligentes des Emotions*. Revue en Intelligence Artificielle RIA édition spéciale Interaction Emotionnelle, vol. 20, N. 4-5, p. 607-620, 2006.
- [13] T.-H.-H. Dang, S. Letellier-Zarshenas, D. Duhaut. GRACE – GENERIC ROBOTIC ARCHITECTURE TO CREATE EMOTIONS. In 11th International Conference on Climbing and Walking Robots and the Support Technologies for Mobile Machines – CLAWAR 2008. Coimbra, Portugal, Septembre 2008
- [14] <http://www.redpsy.com/guide/>, 2006
- [15] <http://www.myersbriggs.org/>
- [16] <http://www.soundsintosyllables.com/>
- [17] <http://www.criff.fr>
- [18] <http://www.fit-pc.com>
- [19] <http://www.valoriam.univ-ubs.fr/emotirob/>
- [20] <http://www.robotis.com>
- [21] <http://www.ftdichip.com/>