

Improving the Expressiveness of Mobile Robots

L. Mayor, B. Jensen, A. Lorotte, Prof. R. Siegwart

*Autonomous Systems Lab, Swiss Institute of Technology, Lausanne, Switzerland
laetitia.mayor@epfl.ch, bjoern.jensen@epfl.ch, antoine.lorotte@epfl.ch, roland.siegwart@epfl.ch*

Abstract

RoboX is a tour guide robot, designed to interact with visitors at Expo02, the Swiss National Exhibition. During five months, ten RoboXs will be in contact with people not trained to autonomous robots and will guide them throughout an exhibition about robotics. The work exposed here was dedicated to make RoboX lifelike and credible in implementing an expression generator. This generator uses the robot capacities in order to best express its internal state which is influenced by its sensor entries and the scenario of the guided tour. In order to validate our approach, a short experiment was set up in order to present results that include appreciations of people not used to robots.

1. Introduction

Our research focuses on a new generation of mobile robots dedicated to operate in public spaces such as information kiosks, receptions, or museums. These robots are designed to communicate with non trained people to give them significant information. In short, we aim to develop robots able to communicate as naturally as possible with human beings. Even if this arises a lot of interesting problems of ergonomics, we will limit ourselves to expressiveness in this paper. Indeed, we believe that synthesized speech is not enough to attract people and keep their interest alive. In fact, it is well known that human-human interaction is not only verbal. When talking with each others, we use voice modulation to enhance our speech. In addition, we use a large number of facial expressions and body movements. Those expressive signs and gestures have been largely studied by psychologists who refer to them as “body language”. In fact other researchers even made the hypothesis that in order to socially interact with humans, robots must be believable and life-like, must have behavioral consistency and have ways of expressing their internal states. The platform “Kismet” was developed to illustrate this theory with voice and facial expressions in the way of infant-like interaction with a caregiver. In the case of Kismet, a complete motivational system was implemented to drive the robot’s action [4]. In our case, the robot is supposed to give tours such as a human guide will give them. The robot actions must therefore be programmed in a scenario, which represents a major difference between Kismet and our robot, RoboX.

Also, previous work with tour guide robots such as Sage [6] or Minerva [5] brought some interesting results about interaction of museum visitors with tour guide robots. For example, visitors tend to take more interest in robots

that look friendly and show a fair face expression. Another example shows the necessity to give the tour guide robots some capacity to react to their environment, especially when disturbed by visitors. Indeed, some people show unfair acts toward the robots, trying to test them. To avoid this kind of problem, the robot should first show that it is aware of the disturbance, and then react properly in an authoritative, thus expressive way.

This paper will then present our strategy for developing a **credible** robot dedicated to guided tours. With credible, we mean that the robot must show that it is aware of its environment and that it is able to react to it in a right way, according to its internal state. Thus, there are three main questions addressed here:

- How to transform the sensor inputs into the internal space?
- How to represent and update the internal state?
- How to express the robot’s internal state?

2. Description of the Robot

RoboX was designed as an exhibition robot. It is working with two main computers:

- a PC, running with Microsoft Windows 2000, that manages the high level tasks, mainly those of interaction and execution of the scenario [1].
- a Power PC, running the real time operating system XO/2, that is devoted to navigation [2].



Figure 1: RoboX, our tour guide robot.

RoboX is able to express its mood with different components. It has a face with two eyes, eyelids and eyebrows overall representing seven degrees of freedom. A matrix of LEDs is mounted on its right eye and different figures can be shown that way. And, some parameters of the

synthesized speech are adjustable.

RoboX is also equipped with different sensor systems. A camera and two laser scanners give the robot a sense of the people around him. The face tracking system is aware of the number of faces in the field of view of the camera and how long they stay. The laser scanners are used to determine how many visitors are all around the robot. At all time, the robot can also detect if someone or something is touching its four buttons dedicated to interaction or its lower body parts, by means of the security switches. Finally, the battery level is known and affect the robot when it is to low.

All the robot's parts are managed by an independent software module implemented as a DLL. The main robot's program parse and execute a scenario file by sending commands to the respective modules.

3. The Expression Generator

First, looking closer at the concept of expressiveness, we can distinguish two main cases in which human beings use the expressions. The first case happens when one wants to emphasize or illustrate its speech. In the second case, the expressions are more like the mirror of the subject's emotions. Since our robot is dedicated to give guided tours, its actions will be predefined by a scenario. We will therefore leave the text illustrations to the scenarist and focus on the second type of expressions that are induced by the internal state of the robot. Nevertheless, it is possible for the scenarist to force the internal state of the robot in order to show a specific expression if needed by the text.

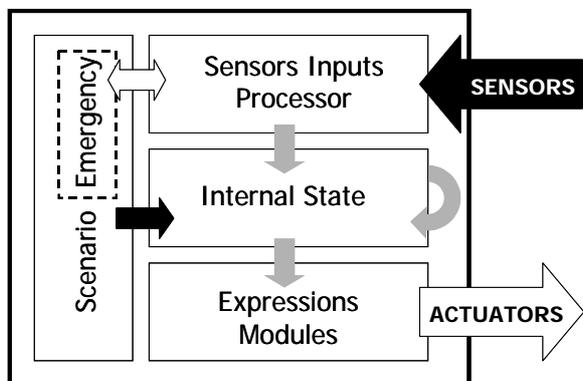


Figure 2: Overall expressions generator process from sensors entries to actuators.

In fact, the robot internal state is even the center of the expression generator we implemented on RoboX. The overall system is represented in figure 2. The generator transforms the inputs from the sensors and the scenario into parameters for all the robot output modules that can show some expressiveness.

For the internal state representation two models seemed interesting: the circumplex affect model of Russel [8] and the Arousal-Valence-Stance affect space [7]. We finally chose the AVS because of its three dimensional representation which is very intuitive to use. Furthermore this representation was previously and successfully

implemented in Kismet.

The robot current state is therefore defined as a point in the three dimensional AVS space (see figure 3). In this space, six expression regions are defined as: sadness, disgust, joy, anger, surprise and fear. These are considered by some psychologists as the six basic expressions [7]. Also, we use the origin of the space as a reference expression that can be considered as a calm state. Of course, other expression regions can be defined in this space. But, we decided to limit ourselves to those seven regions in order not to overwhelm the visitor with many reactions to subtle for our robot expressive capacities.

We will now present the expressive modules and the sensor inputs of RoboX. Then we will expose more details about the internal state.

4. The Expressive Modules

Expressiveness is an old field of investigation. Darwin for example was already interested in the matter and did study its evolution among the species. In fact, expressiveness is not only a human characteristic. Animals are expressive in their own ways, which as human we did learn to recognize. But our robot is neither human-like nor animal like, its appearance is definitely mechanical and rather asymmetric. The visitors that will have to interact with RoboX will discover him and have about fifteen minutes to spend with him. They will not have the time to learn how the robot express itself. The main problem here is therefore: How can such a mechanical system show expressions that are intuitive enough to be understood by anyone at first sight?

Our strategy is to use all the hardware possibilities of the robot in order to show complete voice and body expressions. Considering the expressions themselves, we chose to mimic human ones and even to exaggerate them as far as the robot's hardware is able to.

4.1 Face Movements

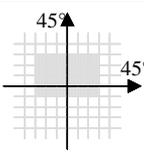
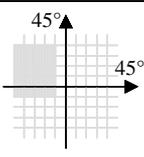
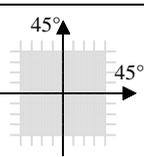
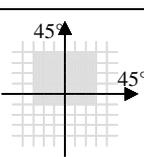
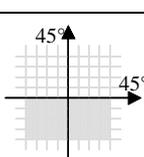
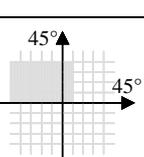
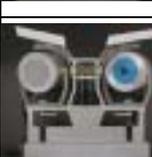
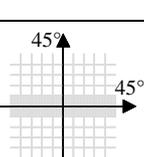
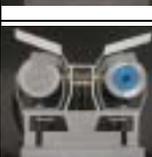
To gain in credibility and to give the most choice possible to the scenarist, three layers of expressiveness were implemented.

- Default behavior: This layer only affect the eyelids and eyebrows. The eyelids blink randomly about every four seconds. The eyebrows position is changed according to the robot current internal state.
- Random movements: On the default behavior, some random movements can be added. Those affect the gaze direction and speed of movement in function of the robot's mood. Indeed, it seems that the gaze direction tells a lot about the state of mind of human beings [10]. We therefore determined a specific window in the eye space, in which the random movement may happen for each expression (see table 1).
- Random sequences: For each basic expression, a sequence of movement using the eyelids, eyebrows and eyes was implemented. These sequences may appear at a specific moment when the robot enters a new mood.

4.2 LEDs matrix

RoboX right eye is mounted with a matrix of 69 blue LEDs. Any figure can be shown with this matrix which works as a little simple screen. The expressiveness of the face can therefore be enhanced using different types of pictures or by changing the light intensity of the LEDs. For example if the light intensity is low, the robot seems sad and tired.

Table 1: Expressive parameters for the face.

	Speed of Movements	Random Window	Main Appearance
Calm	normal		
Fear	slow		
Surprise	very fast		
Joy	fast		
Sorrow	very slow		
Disgust	normal		
Anger	fast		

As for the eye movements, the LEDs expressiveness was implemented in two layers. The first layer corresponds to a default picture which shows an iris. Its pupil size is determined by the robot's mood. The idea behind this iris is to gain some symmetry since the left eye with the camera has a blue iris too. The parameters we choose for the LEDs matrix are given in table 2.

For the second layer six pictures symbolizing the six basic expressions were designed. They appear at the

same time that the random sequences of the eyes movements to enhance them.

Table 2: Expressive parameters for the LEDs matrix.

	Pupil opening diameter
Fear	Medium
Surprise	Large
Joy	Large
Sorrow	Medium
Disgust	Small
Anger	Small

4.3 Voice

Speech synthesis was used for the robot. One of its advantage is that some parameters, such as pitch, volume and rate can be changed while the robot is speaking. We therefore used those parameters in order to add expressiveness to the robot's voice. The following table shows how the parameters are modified for the six basic expressions. The parameters values are described relatively to the "calm" expression that we take as reference.

Table 3: Expressive parameters for the Voice

	Pitch	Rate	Volume
Fear	high	very fast	medium
Surprise	very high	very fast	very loud
Joy	high	fast	loud
Sorrow	little low	slow	very soft
Disgust	low	very slow	soft
Anger	very low	very slow	very loud

5. What affects the robot?

The internal state of RoboX can be modified by two different entries: scenario and sensor inputs. In fact, according to the scenario, the robot may show specific expressions at specific moment in its talk. The scenarist has therefore a simple tool to influence the robot internal state when it is desired. Those programmed expressions can easily be included in the robot's talk, using the software developed for creating scenarios for the robot in which a "behavior task" is available. The scenarist may select one of the seven basic expressions, or he may choose any point in the AVS space. Then he can decide how long each expression should affect the robot.

But when no expression is specified by the scenario or when no scenario is running, the robot internal state is a function of the sensor inputs. In order to determine how to map the sensor inputs space into the AVS space, we defined the robot goals and needs.

In fact, RoboX was designed for two main purposes: navigation and interaction. Its expression should therefore reflects those two goals and help him achieving them. For example, if some visitor is blocking RoboX's path, the robot should be able to show its unhappiness. Also if the robot is alone and seeking for visitors, it

shows joy as soon as it detects someone in its perimeter.

Table 4: Sensor inputs and corresponding affects

	Signal Type	Affect
Face Tracking	nobody in sight	sorrow
	< X persons in sight	joy
	> X persons in sight	disgust
People Tracking	nobody in sight	sorrow
	< X persons in sight	joy
	> X persons in sight	disgust
Buttons	touched without question asked	anger
Scanner Lasers	blocked in front/back	anger
Bumpers	touched in front/back	anger
Battery	low level	sorrow

Aside its goals, RoboX has two main needs. The first is battery level. If the batteries are running low, the robot is unable to continue giving tours and need to go to a docking station. The second need is dictated by the robot's job itself. To insure a specified visitor flow, a tour given by RoboX should not be more than fifteen minutes long. So the robot expressions must help him finishing its tour in time. Events that could delay the robot in its tours are of different kinds, but most of them are linked with the behavior of the visitors. We define three types of disturbances:

- Misuse of the interactive buttons: Happens when a visitor is playing intensively with the robot buttons, even when the robot is not waiting for an answer. In this case, the robot is not directly troubled. But the other visitors may not be able to follow the robot's tour with someone staying in front of the guide and playing with its buttons.
- Bumpers activated: When something touches the robot's low body part, its motion system is stopped and the robot cannot move anymore as long as the bumper is not released.
- Path blocked: The robot's navigation is completed by an obstacle avoidance system. If an object is in the robot's path, the robot stops and try to figure out another path. In normal conditions, this path recomputation is not extremely time consuming. But if this happens very often or if someone is deliberately blocking the robot's path, it may really delay the robot.

In all this cases, RoboX's mood tends toward disgust or anger. And if the disturbances happen to often, what we call emergency scenarios may be launched.

It may also happen that more than one signal is detected at the same time. In this case, the sensor that is in the lowest position in table 4 is taken into account.

5.1 Emergency scenarios

The idea of the emergency scenarios is to show to the visitor that the robot is aware of its environment and that it is able to react to it, especially when it is disturbed. We hope to increase the robot's credibility by this mean. There are in fact six cases in which emergency scenarios can be launched: the batteries are running low, someone is playing with the robot buttons, the robot's path is

blocked forward or backward, and the front or the back bumpers are touched.

When such a scenario is triggered, the robot main program dynamically change the main scenario in order to include an emergency sequence in it. The robot may then excuse itself if its battery level is to low, or it may ask the visitors to let it do its job without disturbing it in the other cases. It is possible to load four different scenarios for each of the six emergency types. And when an emergency is detected, one of the four available sequences is randomly chosen.

Of course, such scenarios are not started every time an emergency signal is sensed, but rather after a certain number of signals of the same type happening in a short period of time (about one minute). These are the conditions we foresee for a visitor trying to really disturb the robot.

6. Internal State of the Robot

RoboX internal state can be described as point in a tridimensional space called the AVS space (Arousal-Valence-Stance) [7].

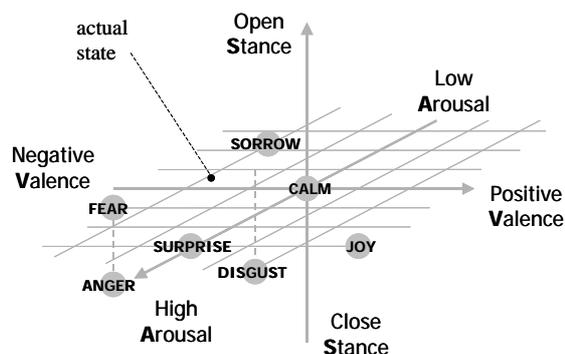


Figure 3: Representation of the six basic expressions in the AVS Space.

As explained above, the internal state of the robot can be modified by sensor entries or by the scenario. Those two inputs signals are transformed in a point of the AVS space according to look-up tables. Using this input point, a new point corresponding to the robot current affect is computed following equation (1).

The T parameter determines the speed of the change of expression. It is chosen in function of the position of the input affect point, more precisely of its arousal coefficient. The reason of that choice is due to the fact that all expressions do not occur at the same speed. For example, surprise is usually instantaneous, in contrary of sorrow which comes slower.

Once the new current affect processed, all the parameters of the expressive modules are updated using equation (2). This is, in fact, equivalent to a weighted mean, using the distance of the current state to the seven predetermined regions in the AVS space as weight. Thus, the more the current state is close to a region, the more the current expression parameters will be influenced by the

parameters of this nearest region.

$$\vec{A}_{new} = \frac{((TA_{old}) + A_{input})}{T + 1} \quad (1)$$

A= point in the AVS space
T= speed of expression change

$$P_{new,n} = \frac{\left(\sum_a \left(\frac{P_{a,n}}{d_{new,a}}\right)\right)}{\sum_a \frac{1}{d_{new,a}}} \quad (2)$$

n = parameter type (eye speed, voice pitch, ...).
a = expression type (calm, fear, joy, ...).
 $P_{new,n}$ = new parameter of type n.
 $P_{a,n}$ = predefined parameter of type n for the expression a.
 $d_{new,a}$ = distance from the new affect point to the region of the expression of type a in the affect space.

This implementation has finally two important advantages:

- The expression space of the robot is linear, which means that wherever the current state point is in the affect space, the robot shows an expression defined by its exact position. This induces that transition between expressions are insured by the system itself.
- The computation is really fast and needs very little CPU time, which means that the robot reactions happen in real time.

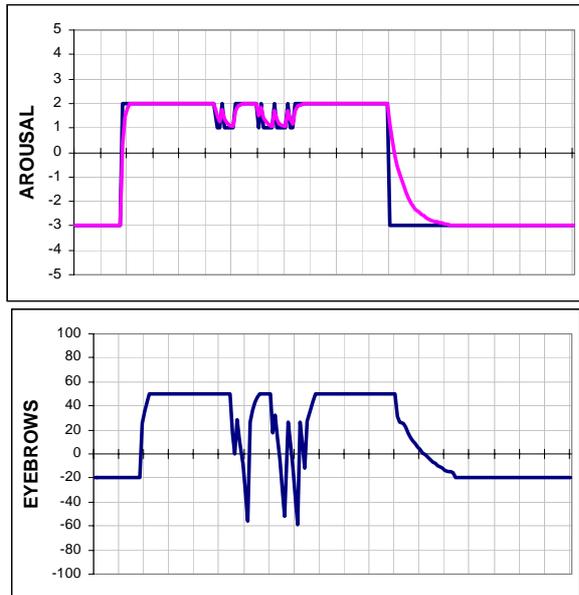


Figure 4: top: Input arousal value (dark) and current arousal value (light); bottom: current eyebrows position.

Figure 4 illustrates this advantages. It shows the arousal value and the corresponding eyebrows position during a short experiment. The robot was running no scenario and nobody was in sight in the beginning. The robot thus showed sorrow until someone arrived. At this time, the arousal value did rise very fast following closely the input arousal signal. The visitor did then try to play with the robot buttons, without being asked to use them. The robot went nervous, and began to lower its eyebrows in a human fashion. As soon as the visitor stopped using the

buttons, the joy expression was triggered. Finally, the visitor left the robots, and it went back to a sad expression.

7. Experiment

We did set up this experiment mainly to test the perception of the expressions (are they intuitive enough to be well understood by anyone?). We therefore prepared three short scenarios of about 30 seconds that we did show twice to our visitors, once with the expression generator running and once without (see table 5 for the exact methodology). The three have the same base, the robot introduces a subject, asks a question about it, waits for an answer (with the buttons) and reacts to the answer. For example, in the “RoboX’s Riddle” theme the robot begins by saying that he likes riddles and then ask the visitors whether he knows “What is small, green and goes up and down?”. RoboX asks then the visitor to touch its red button if he knows the answer, or the green button if he doesn’t know the answer. Then in function of the visitor’s choice, RoboX will give the riddle answer (“A little pea in a lift!”), or it will comment that “Everybody knows this riddle already. Only robots find it funny anymore”. All of the three themes contained some facial expressions that were only text illustrations. Since those are not affected by the expression generator, they did appear every time. Also, for practical purposes we only used the face and voice modulation, and not the LEDs and robot movements.

Table 5: Experiment methodology

	Type	Expressive
Scenario 1	RoboX’s riddle	Yes
Scenario 2	What is my name?	No
Scenario 3	RoboX’s riddle	No
Scenario 4	Today’s weather	Yes
Scenario 5	What is my name	Yes
Scenario 6	Today’s weather	No

We then had 37 persons who took part to the experiment. They were asked to answer three questions after having seen each scenario about how well they understood the robot, how expressive it was and what type of expressions it just performed (they had the choice between: none, fear, anger, joy, surprise, sorrow, disgust or other). Finally, we asked the visitors to answer three more general questions to get some more precise idea of what they liked and did not like in RoboX.

The results show that overall, the visitors understood well the robot’s synthesized speech (figure 5). Nobody checked the “very badly” possibility which is very encouraging. Nevertheless, the visitors did need some time to get accustomed to RoboX’s voice. And, while trying to concentrate on the voice, they forgot to watch at the face expressions. But it seems that one scenario is enough to get used to RoboX face and voice. About one third of the visitors made the comment that they had some difficulties understanding the scenarios 2 and 5 which corresponds to the “What is my name?” theme. In these scenarios the robot talks about its name and there-

fore uses the “RoboX” word which is not in the common language.

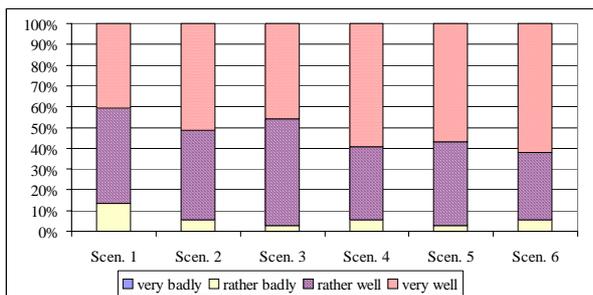


Figure 5: Comprehension for the six scenarios.

We also wanted to know if the visitors were able to make some difference between the scenarios played with the expression generator running (scenarios 1, 4 and 5) and the one without the expression generator (scenarios 2,3 and 6). The results are shown in figure 6. One can observe that the difference for the “Robot humor” (scenarios 1 and 3) and “Today’s weather” (scenarios 4 and 6) themes are significant. The “What is my name?” theme (scenarios 2 and 5) was overall implemented to be less expressive than the others, which implies fewer text illustrations and fewer behavioral expressions. The visitors did notice this lack of expressions too.

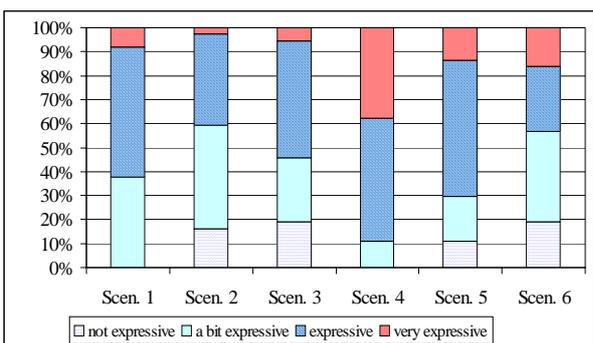


Figure 6: Expressiveness for the six scenarios.

The last question was whether the implementation of the expressions was understandable enough or not. It turned out (figure 7) that some expressions were much more recognized than other. Especially, sorrow, joy and fear were well recognized. But, nobody did see disgust at the right time. Also, anger and surprise were poorly noticed. This arises the question about adding a mouth to RoboX because among human beings, the disgust and surprise expressions use intensively the mouth. Also, the lack of a mouth was mentioned explicitly 11 times in the answers to the general questions. Anyway, the experiment was pursued without the use of the LEDs matrix, which may fill this lack.

Finally, the theme “Today’s weather” (scenarios 4 and 6) was mentioned 9 times by the visitor as the best theme. This success is probably due to two main factors. First, the expressions that appear are fear, joy and anger, which are the one that the visitors did recognize best. Second, in this scenario, the robot tries to look out the window, asks the visitor what the weather is since it cannot look out by itself. It therefore shows some simple intention,

and act in order to fulfill it.

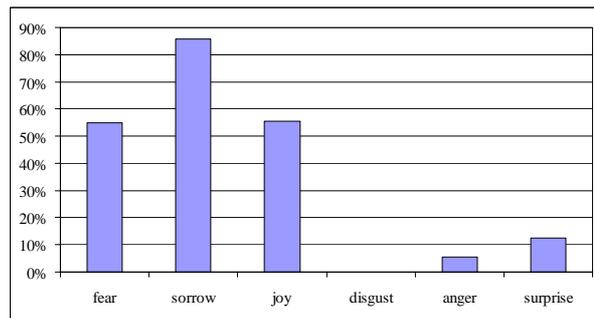


Figure 7: Recognition of the six basic expressions

8. Conclusions

This paper went through the implementation of an expression generator dedicated to a tour guide robot. In order to add credibility to our robot, we used all the modules that could show some expressiveness: face, voice and body movements, and emergency scenarios were added to the system. The system was then tested with potential visitors. The results show that the peoples that discovered RoboX for the first time need only a short time to get used to it. The results also showed that the lack of an expressive mouth may induce problems for the expressions recognition. Finally, the scenario were RoboX showed intentions and much expressions seemed to be preferred by the visitors which tends to prove that for a good human robot interaction, the robot should be able to show behavioral consistency and have ways of expressing its internal state.

9. References

- [1] B. Jensen et al., “Interactive Mobile Robotics: RoboX, a case study.”, *to be published*.
- [2] K.O. Arras et al., “Multisensor On-the-Fly Localization: Precision and Reliability for Applications”, *Robotics and Autonomous Systems*, 34(2-3), 2001.
- [3] C. Breazeal and B. Scasselatti, “How to built robots that make friends and influence people”, *Proceedings of the IEEE International Conference on Intelligent Robots and Systems*, Kyonjiu, Korea, 1999
- [4] C. Breazeal, “A Motivational System for Regulating Human-Robot Interaction”, *Proceedings of American Association for Artificial Intelligence*, Madison, WI, 1998, pp.54-61
- [5] S. Thrun et al., “MINERVA: A Second-Generation Museum Tour-Guide Robot”, *Proceedings of IEEE International Conference on Robotics and Automation*, 1999
- [6] I. Nourbaksh et al., “An Affective Mobile Robot Educator with Full-time Job”, *Artificial Intelligence 114 (1-2)*, October 1999, pp. 95-124.
- [7] P. Ekman and R. Davidson, *The Nature of Emotion: Fundamental Questions*, Oxford University Press, New York, 1994.
- [8] J.A. Russel, “A Circumplex Model of Affect”, *Journal of Personality and Social Psychology* 39, 1980, 1161-1178.
- [9] M.-A. Descamps, *Le langage du corps*, Presses Universitaires de France, 1989
- [10] R. Dantzer, *Que sais-je?: Les émotions*, Presses Universitaires de France, 1988