Seeing Signs: Pre-Adolescent Recognition of Social Body Language in Humanoid Robots

Karen C. Aragon
CUNY Brooklyn College, 2900 Bedford Avenue, Brooklyn, NY

Abstract

The utilization of humanoid robots, such as Aldebaran's NAO, is expanding in scope, most prominently, as companions in both therapeutic and educational environments in order to improve the quality of human life. This paper focuses on emulating emotion and expressive behaviors in an anthropomorphic robot in order to foster meaningful and intuitive humanoid-robot interaction, specifically with pre-adolescent children. We utilize non-verbal, full body movements to model the six basic emotions and incorporate these emotions into an interactive and robust educational game. We present the design and experimental evaluation of the emotionally expressive humanoid robot during a socially interactive game of 'Charades'. We also lay the groundwork for an educational game that enables a child user to interact one on one with the robot that incorporates various modalities including the emotions previously mentioned. We then propose future methods and ideas to improve the human-robot interaction in the proposed educational game.

Keywords: Human-Robot Interaction, Emotions , Non-verbal Body Expression, Education.

1 Introduction

The interest in advancing humanoid-robot interaction stems from work in the HRI community that focuses on fostering meaningful and intuitive communication between robots and people. This entails modeling aspects of human-human interaction in order for people to perceive the robot as "real," that is, convincingly emotive and social.

One aspect of human-human interaction is a person's ability to detect emotional cues of another person during conversation [1, 7]. A person smiling, raising their arms, and extending their fingers might indicate joy and happiness, while a person with hunched shoulders and a lowered head may indicate a state of sadness.

Furthermore, research suggests that incorporating realistic emotional displays into a computer simulated character can have the general effect of making it seem more believable and human-like, and thereby cue the user to interact with the character as if they were interacting with another person [7,10].

In this paper we presume that humans can project this same schemata of emotional-visual cue-seeking onto the body language of a robot. By giving humanoid robots the ability to exhibit appropriate emotions during communication, humans can then gain some insight to the internal state of the robot itself. We assume that humans will anthropomorphize humanoid robots once they create meaningful engagements.

1.1 Emotional Anthropomorphic Robot

We describe an anthropomorphic robot, the NAO, and how we endow it with the ability to communicate human-like emotions through full body movements. The ability to program a set of body languages into the NAO suggests a series of correlating emotional states. We aim is to determine which positions most successfully allow for more engaging and meaningful interactions between a robot and a human, and to thereby improve human-robot interaction.

1.2 Widening Scope

The widening scope of use for robots extends to the entertainment, education, domestic, and clinical sectors, among others. Growing attention is given to socially-assistive robots, from ones that give care to individuals with cognitive disabilities [8,12,13], robots that assist in physical rehabilitation [8], and companion/therapeutic robots that give special care and attention to the elderly and children with extended hospital stays [8,13]. Examples of these robots are further explained in Section 2.

In general, robots used in these fields are used as companions and looked to for supportive help rather than as machines with limited abilities. As researchers continue to endow humanoid robots with various aspects of human-human interaction such as expressing emotions, maintaining appropriate eye contact and gestures, and an expressive voice, this ultimately can lead to an even wider acceptance of robots and dispel peoples' perception of robots as merely mechanical tools.

1.3 Emotional Expression during Game State

We are also interested in the prospect of using humanoid robots as an educational medium. We first address the research issue of whether full body emotional expression through the robot NAO can be identified by pre-adolescent children, and then we lay the foundation for further research into whether these emotions facilitate a child's ability to learn in a teacher-student scenario.

Our aim with developing an interactive game is to create a comfortable learning environment for children to have an engaging and educational experience with an anthropomorphic robot as teacher. During the game, the NAO robot uses human-like traits such as encouragement and positive feedback to promote student learning and motivation.

2 Related Work

There are many studies that investigate and develop anthropomorphic robots designed to interact and assist their users using emotional-visual expressions. Their work aims to bring attention to robots that are treated as companions rather than as tools.

Gratch and Marella's paper focuses on the use of emotion in education settings and specifically addresses the function of emotion in a computational system. They investigate ways the emotion functions can be modeled and how the emotions externally manifest to a user [7]. They describe two different camps of emotion functions, *intra-organism*, which views emotion as an entity mediating mental processes and *inter-organism*, which views emotion as an entity that impacts social interaction. The paper also presents two approaches to implement the emotion functions, *Communication-driven* and *Simulation-based. Communication-driven* approaches intentionally convey emotion by utilizing verbal and nonverbal forms of communication. In *Simulation-based* approaches, "emotion displays are dependent on the agent's simulated emotional state which reflect something about the current state of information processing of the agent [7]." It is important to note that we utilize the *Communication-driven* approach which we further explain in Section 3.4.

Studies suggest the ability for humanoid robots to convey appropriate emotions through body postures. Using the Laban movement framework researchers aim for a standardization of expressing emotions through body movements [1, 5].

At the Artificial Intelligence Lab at MIT, Kismet was designed with the emotional level of a human infant [2]. Modeling behaviors from research on human emotion, Kismet pines for attention when ignored, emulates happiness when a user engages in close proximity, and feels agitated when overstimulated. Kismet uses various emotional expressions to engage and learn from users.

While Kismet only utilizes facial expressions, further studies incorporate full body movements to emulate emotions in socially-assistive robots [1,5,6, 14]. Pelachaud et al present a humanoid robot that displays expressive gestures during storytelling [5].

Incorporation of emotional expressions in the robot aim to maintain human users' attention and interest during social engagement.

Results from a study in Korea, found that home robots were superior in promoting and improving students concentration, interest, and academic achievement when compared to other learning programs such as web-based instruction [4].

3 Design and Implementation of 'Charades' game

3.1 Approach

Our approach models four of the the six basic emotions: anger, fear, joy, sadness. using full body motions and LED colors in the eyes. No other external clues such as sound or voice were incorporated. We incorporate these expressive emotions into a game of 'Charades' which enables a user to interact with the robot itself.

3.2 Software Specifications

Choregraphe, version 1.12.5, provided by Aldebaran, was used to create the kinesthetic movements in the NAO robot. Choregraphe is a GUI-based, multi-platform desktop application. that allows for intuitive manipulation and control of the NAO robot [3].

To create complex behaviors such as a choreographed dance, walking to a destination, or tracking an object, a user connects a series of boxes containing code to run simultaneously or sequentially on the NAO robot. An example of the Choregraphe interface can be found in Section 6.4. Each box contains a corresponding Script box written in Python.

The default Script box, shown below, defines a class named MyClass which inherits from GeneratedClass which provides the built-in functions: onLoad(), onUnload(), onInput_onStart(), onInput_onStop() [3]. The script box also allows you to add or import your own Python functions.

```
class MyClass(GeneratedClass):
        def __init__(self):
                GeneratedClass.__init__(self)
        def onLoad(self):
                pass
        def onUnload(self):
                pass
        def onInput_onStart(self):
                pass
        def onInput_onStop(self):
                self.onUnload()
                passs
```

In order to create the movements for the different emotions, we stored joint positions for the head, arms, legs, and digits using key frames.

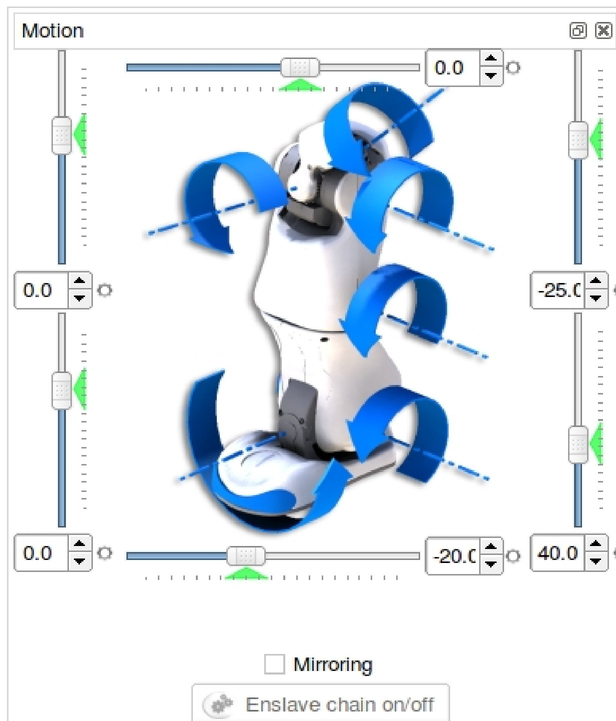


Figure 1 Manipulation of joint angles of right leg

3.3 Hardware Specifications

Aldebaran's NAO robot was the original creation of students at the French design and management school, Creapole. Their aim was to create a robot with affordability, performance, modularity, and open architecture.

It is boasted as the first fully programmable and emotional robot and replaced Sony's AIBO as the robot used for Robot Soccer World Cup. It is equipped with two CMOS cameras, four microphones, two speakers, laser, sonar, and bump sensors. These features allow for facial, object, and speech recognition, text to speech synthesis, dynamic movements, amongst various other abilities.

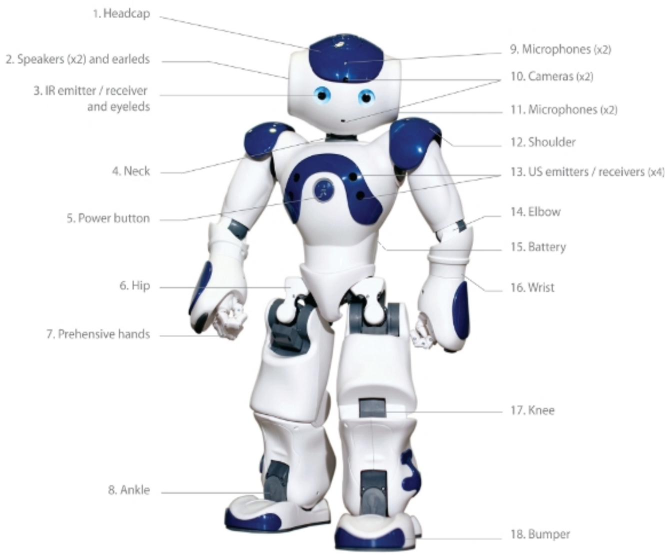


Figure 1: Features of the Aldebaran NAO robot

The NAO robot lacks the motor ability to display emotions in the face, but has 25 degrees of freedom to display emotions through full body movements. Included in the package is an intuitive GUI based software program called, Choregraphe, which is explained in Section 3.2.

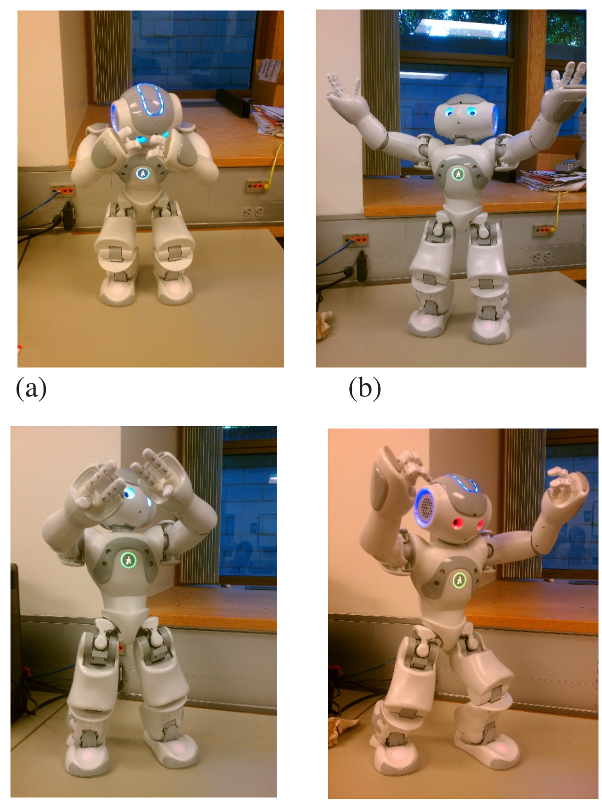
3.4 Framework for Modeling Emotions

A game of 'Charades' was utilized as an effective and fun means to engage the children to identify the specific emotions of the robot.

The table below lists characteristics of each of the programmed emotions.

Emotion	**Characteristics**
Anger	red eyes, erratic positioning of arms, clenched fists, jerky movements
Sadness	Blue eyes, hunched shoulders, lowered head, shaking of head from side to side, hand covering face
Joy	Open arms, raised head, chest out, digits open and fully extended
Fear	hands positioned in front of body, cowering movements, crouching down, shrinking

Only a short sequence of movements was needed to program the emotions onto the NAO. Facial expressions were not used on the robot because it lacks actuators in the face to simulate expressions



(a) (b)

(c) (d)

Figure 4 a) sadness b) joy c) fear e) anger

Color was utilized though LED in the NAO's eyes only with fear and sadness. Red LED was used to indicate anger and Blue LED was used to indicate sadness.

4 Experiments and Results of 'Charades' game

The goal of the experiment, using the 'Charades' game is to measure the ability of children to identify specific emotions using nonverbal full body postures. In order to determine whether the emotions are representative of the affective states they were designed to portray, we played a game of 'Charades' where children were instructed to identify the emotion displayed by the robot.

4.1 Test Group

A group of 77 children were used for this experiment. We conducted the experiment with eight groups of ten children. The children were part of a summer robotics camp held by the University of Minnesota. The children were placed in a half circle around the NAO robot.

4.2. Experiments

Prior to the experiment, the children were each given a questionnaire where subjects could circle one of four emotions: joy, anger, fear, or sadness. During the experiment, the NAO robot first gives an introduction then instructs the students that the robot will display a series of eight different emotions. The robot explains that task for the children is to identify and circle the appropriate emotion listed on the questionnaire. We further instructed the children that questionnaires were to be completed independently, that no answers were to be given aloud, and that there were no right or wrong answers. The NAO performed 2 variations of anger, joy, fear, and sadness. Then questionnaires were collected.

4 Results

The results of our questionnaire show that pre-adolescent children could correctly identify the appropriate display of fear, anger, sadness, or joy.

On average, between the two displays for each emotion, 82% correctly identified Joy, 74%, correctly identified Sadness, 91% correctly identified Fear, and 94% correctly identified Anger.

The results further indicate that of the two different displays for each emotion, one was more identifiable than the other. Joy had the widest gap with 67% of the participants identifying Position #1 (first joy) correctly and 97% correct for Position #8 (second joy). There was also a disparity between the two variations of Sadness with 58% correctly identifying Position #2 (first sadness) and 89% correctly identifying Position #5 (second sadness).

Various factors such as position in the sequence of behaviors, ambiguous movements, or inappropriate LED color could have attributed to the differences in results between both variations of Joy and Sadness.

The recognition rates for anger were consistently high – 92% and 90%, respectively.

	Joy	Anger	Sadness	Fear	**% Correct**
#1 Joy	52	9	8	8	**67**
#2 Sadness	1	2	45	29	**58**
#3 Anger	0	71	5	1	**92**
#4 Fear	0	1	6	70	**90**
#5 Sadness	0	1	69	7	**89**
#6 Anger	4	70	2	1	**90**
#7 Fear	12	1	6	58	**75**
#8 Joy	75	1	3	0	**97**

Table 1 Summarizes the data results from the experiments with 77 participants

5 Conclusions

Overall, the data gathered from the experiments suggest that pre-adolescent children, with minimal experience with a NAO robot, are able to identify appropriate emotional displays emulated on the robot itself. This section described a design and implementation of nonverbal emotional behaviors on a humanoid robot.

The results of the 'Charades' game will be used to better guide and inform the design and implementation of future research into human-robot interaction.

6 Design and Implementation of Interactive Educational Game

The aim of the educational game is to incorporate various modalities such as object and facial recognition, text to speech synthesis, speech recognition, and emotions previously mentioned, to facilitate meaningful interaction between human and humanoid robot. Indeed, some education researchers have argued that such nonverbal displays can have a significant impact on student intrinsic motivation (Lepper, 1988).

The design of the game follows an identification schema where the robot takes on a teacher role asking the user (student) to search and identify objects. A human is needed to set up the game and help the user if instructions are not clear.

6.1 Software and Implementation

Both the Choregraphe and the NAOqi API in Python were utilized to develop the game. Motions, such as full body gestures to emulate emotions, changing LED color of the eyes, and incorporating music were programmed through Choregraphe. We were able to 'teach' the robot to recognize various objects using the video monitor in Choregraphe, which we discuss further in Section 6.2.

Using the NAOqi API, provided by Aldebaran, further modalities such as speech recognition and text to speech synthesis were incorporated into the game. We utilized the *ALTextToSpeech* module, provided by the API, to enable to NAO to speak. The *ALSpeechRecognition* module was used to enable the robot to recognize words and phrases spoken by the human user. The Python script below, demonstrates how to enable the NAO robot to perform speech recognition. In order to use modules provided by the API, you need to create a proxy onto the module. You then used the subscribe() to subscribe to the module and unsubscribe() to quit.

```
asrProxy = ALProxy('ALSpeechRecognition',
IP,    PORT)
tts = ALProxy('ALTextToSpeech',IP,PORT)

#Sets the words to recognize
wordList= ['yes','no','okay']
asrProxy.setLanguage('English')
asrProxy.setWordListAsVocabulary(wordList)

tts.say('Would you like to play a game?')

#start Speech Recognition
asrProxy.subscribe('MyModule')

WordRecognized= False
ALMemoryKey= 'WordRecognized'
recognizedWord=
       self.memProxy.getData(ALMemoryKey, 0)

#enter speech recognition loop
```

6.2 Learning the objects

The NAO robot first needs to be 'taught' to recognize the objects to be used during game play by utilizing the vision monitor in Choregraphe. After the images are learned and stored in the database of the NAO, the object recognition module should be tested to assure that the robot is able to identify the correct object when shown by the user.

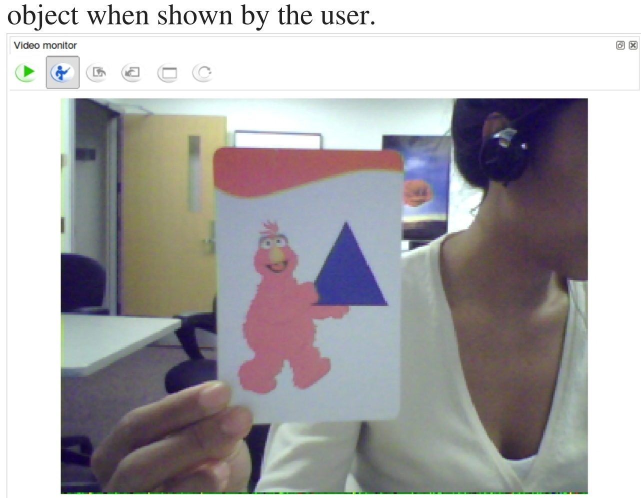


Figure 5 Vision Monitor Module of Choregraphe

Figure 5 shows the vision monitor module in Choregraphe. The 'Learn' button is used to obtain an image of the flashcard using the NAO's cameras. Once the image is captured, the user manually clicks around the perimeter of the flashcard image. The specific image is then stored in NAO's database by selecting the 'Import' button. Once all images are stored into the NAO's database, the NAO will be able to perform object recognition.

6.3 Setup

A series of flashcard are placed in front of the student with the NAO on the table in a sitting position about two feet away. The flashcards could be of any subject matter, but the NAO robot must first be taught the images, explained in Section 6.2. The student is then told by the human conductor to pay attention to the robot's instructions.

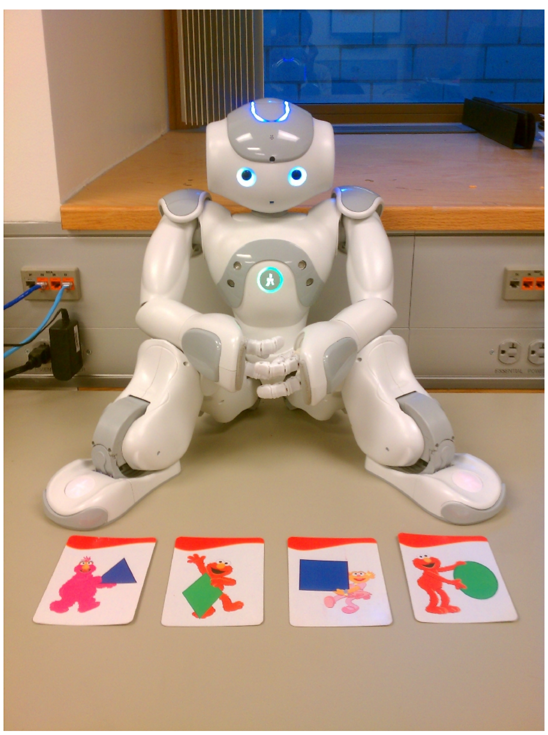


Figure 6 Shows setup of game

6.4 Game play

The NAO gives a short introduction and asks the user if he/she would like to participate in the game. If the user gives a 'yes' response, the NAO proceeds into a series of questions asking the user to identify a specific shape. The NAO waits for the user to identify the correct shape by raising the card to eye level of the robot where it is fully visible for the cameras.

Once the card is shown to the robot, the robot enters the vision-recognition module of Choregraphe where it attempts to match the image with an image in the vision database. If a correct match is found, the NAO elicits a positive response which includes blinking LED lights in the eyes, joy-like body movements, and encouraging and congratulatory words. If the child chooses the incorrect object, the NAO kindly asks the user to try again and re-enters the vision recognition loop. We are careful to not elicit a strong negative attitude, such as anger, since the aim of the game is the game is to facilitate and foster effective learning between humans and robots.

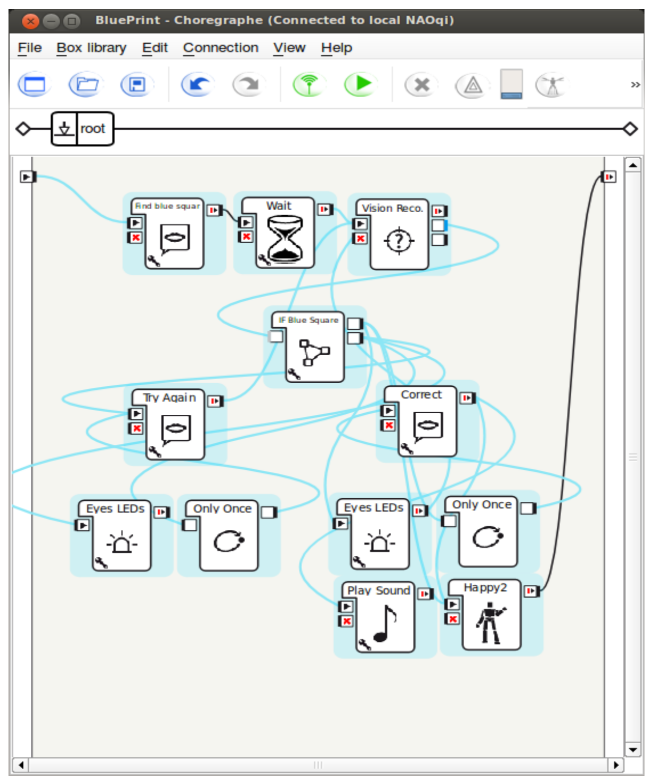


Figure 3: Simple game flow of one question. The Nao robot asks a question, waits for a response, once the object is seen by the cameras it enters the Vision Recognition phase, if the object is correct it enters the correct phase, if

incorrect, the robot asks the user to try again and game play loops back to the Vision Recognition module.

7 Further Work

The general goal of the Interactive Game explained in Section 6 is to create a comfortable educational environment in which a humanoid robot is able to effectively communicate and facilitate learning with a user.

Much interest is given to investigating the manners in which humans themselves react to emotionally expressive humanoid robots. Future work involves enabling the humanoid robot to communicate in either an extroverted or introverted teaching style during the game state. The question of interest is 'To what extent does personality type affect the ability of students to learn?'

Another facet that deserves further research is to enable the robot to change personality according to speech clues in the users voice. If a robot detects that the child is bored, the robot could take on a more extrovert and encouraging personality.

8 Acknowledgments

I would like to thank Professor Maria Gini of the University of Minnesota for her guidance and advice, and DREU (Distributed Research for Undergraduates) for this 10-week research fellowship opportunity.

References

[1] Barakova, E. I., Lourens, T. “Expressing and Interpreting emotional movements in social games with robots.” Personal and Ubiquitous Computing, vol. 14, pp. 457 – 467, 2010.

[2] Breazeal, C., “Emotion and sociable humanoid robot” . Int. J. of HumanComputerStudies, vol. 59, pp.119-155, 2003.

[3] “Choregraphe User Guide.” Aldebaran Robotics. Web. Aug 2012. <http://www.aldebaran-robotics.com/documentation/software/choregraphe/>.

[4] Han, J., Jo, M., Jones, V., Jo, J. “Comparative Study on Educational Use of Home Robots for Children.” Journal of Information Processing Systems, Vol.4, No.4, 2008 .

[5] Le, Q.A., Hanoune, S. and Pelachaud, C. “Design and Implementation of an expressive gesture model for a humaoid robot.” 11th IEEE-RAS International Conference of Humanoid Robots (Humanoids 20122), 2011.

[6] M. Kipp and J.-C. Martin, ‘Gesture and emotion: Can basic gesturalform features discriminate emotions?’, in International Conference on Affective Computing and Intelligent Interaction (ACII-09). IEEE Press, 2009.

[7] Marsella, S., Gratch, J., and Petta, P, “Computational Models of Emotion.” In Scherer, K.R., BÃ¤nziger, T., & Roesch, E. (Eds.) A blueprint for an affectively competent agent: Cross-fertilization between Emotion Psychology, Affective Neuroscience, and Affective Computing. Oxford: Oxford University Press, 2010.

[8] Matarić, M. J., "Socially Assistive Robotics", *IEEE Intelligent Systems*, 21(4), Jul/Aug 2006, 81-83.

[9] “Python NAOqi API .” Aldebaran Robotics. Web. Aug. 2012. <http://www.aldebaran-robotics.com/documentation/ref/python-api.html>.

[10] Reeves, B., Nass, C., 1996. “The Media Equation.” CSLI Publications, Stanford, CA.

[11] Scassellati, B., 2000. “Foundations for a theory of mind for a humanoid robot”. Ph.D. Thesis, MIT Department of Electrical Engineering and Computer Science, Cambridge, MA.

[12] Schmitt, L. M., Diehl, J. J., Villano, M., Crowell, C. R. 2010. “The clinical use of robots for children with Autism Spectrum Disorders: A critical review.”

[13] Wallbott, H.G., ‘Bodily expression of emotion’, European Journal of Social Psychology, 28, 879 – 896, 1998.