Bender: A General-Purpose Social Robot with Human-Robot Interaction Capabilities

Javier Ruiz-del-Solar,

Department of Electrical Engineering, Universidad de Chile

Advanced Mining Technology Center, Universidad de Chile

Mauricio Correa and Rodrigo Verschae

Advanced Mining Technology Center, Universidad de Chile

Fernando Bernuy

Department of Electrical Engineering, Universidad de Chile

Advanced Mining Technology Center, Universidad de Chile

Patricio Loncomilla

Advanced Mining Technology Center, Universidad de Chile

and

Mauricio Mascaró, Romina Riquelme, and Felipe Smith

Department of Electrical Engineering, Universidad de Chile

The main goal of this article is to describe Bender, a general-purpose social robot with human–robot interaction capabilities, and to report and analyze its applicability in three different natural environments: (i) home settings, (ii) school classrooms, and (iii) public places. Evaluation of the robot’s performance is based on its ability to detect and identify humans in those settings, its ability to express emotions, and its ability to interact naturally with humans. The obtained results, although preliminary, indicate that the robot has great acceptance from expert and non-expert human users alike, and that it is able to interact successfully with humans using human-like interaction mechanisms, such as speech and visual cues (particularly facial information). It is remarkable that the robot even has the potential to teach children in a real classroom setting. Further experiments will be conducted in order to better assess the robot’s potential.

Keywords: human–robot interaction, social robots, robot acceptance

1 Introduction

Social robots are gaining increasing interest in the robotics community. A social robot is a subclass of a mobile service robot designed to interact with humans and to behave as an assisting partner, providing entertainment, companionship, and communication interfaces. Its development is based on the expectation that the morphology and dimensions of social robots allow them to operate adequately in human environments. It is projected that social robots will play a fundamental role as companions for elderly people and as entertainment machines in the coming years.

Authors retain copyright and grant the Journal of Human-Robot Interaction right of first publication with the work simultaneously licensed under a Creative Commons Attribution License that allows others to share the work with an acknowledgement of the work's authorship and initial publication in this journal.

Journal of Human-Robot Interaction, Vol. 1, No. 2, 2012, Pages 54-75, DOI 10.5898/JHRI.1.2.Ruiz-del-Solar

Among other capabilities, social robots should be able to

- move in human environments,
- interact with humans using human-like communication mechanisms (speech, body and facial expressions, and hand gestures),
- manipulate objects,
- determine the identity of the human user (e.g., “owner 1,” “unknown user,” “Peter”) and sense his/her mood (e.g., happy, sad, excited) to personalize their services,
- store and reproduce digital multimedia materials (images, videos, music, digitized books),
- connect humans with data or telephone networks,
- collect and report data from different sensor networks and the Internet, and
- be empathic (humans should like and feel understood by them).

In addition,

- their usage should be natural without requiring any complex technical or computational knowledge, and
- they should be robust enough to operate in natural environments.

Social robots with these abilities can assist humans in different environments such as public spaces, hospitals, home settings, and museums. Furthermore, social robots can be used for educational purposes.

Social robots should be capable of being accepted by different types of human users, including non-expert ones such as children and the elderly. We postulate that, in order to gain acceptance, it is far more important to be empathic and to evoke sympathetic responses from humans than to have an elaborate and elegant design. Moreover, to produce effective interaction with humans, and even to enable humans to behave as if they were communicating with peers, it has been suggested that the robot's body should be “based on a human's” (Kanda, Ishiguro, Imai, & Ono, 2004) or be human-like (Hayashi et al., 2007). We propose that it is important to have a somewhat anthropomorphic body, but that having a human-like body is not necessary. Many researchers have also mentioned the importance of the robot maintaining eye contact, i.e., tracking or gazing at the face of the speaker when interacting with humans (Ishiguro, Ono, Imai, & Kanda., 2003; Kanda, Ishiguro, Ono, Imai, & Nakatsu, 2002; Matsusaka et al., 1999; Nakadai, Hidai, Mizoguchi, Okuno, & Kitano, 2001). We agree that these mechanisms for maintaining attention are important for the human user. In particular, face detection of the user allows the robot to keep track of it and face recognition enables the robot to identify the user, to personalize its services and to make the user feel important and emotionally understood (e.g., “Sorry Peter, can you repeat that?”). It is also relevant that interaction with the robot be natural, intuitive, and based primarily on speech and visual cues, as there are individuals who are technologically handicapped.

The question is how to provide all these capabilities. We believe that these capabilities can be achieved if the robot has a simple and anthropomorphic body design, is able to express emotions, and has human-like interaction capabilities, such as speech, and non-verbal communication capacities, such as facial expressions and hand gestures. We also believe that it is important for the cost of a social robot to be reasonable, particularly if our final goal is to introduce it in typical human environments where the users will be normal persons with limited budgets. Taking all the above into consideration, we developed *Bender*, a general-purpose social robot that incorporates these characteristics, at the Robotics Laboratory of the Department of Electrical Engineering of the Universidad de Chile.

The main goal of this article was to describe this robot and to report and analyze its applicability in three different natural environments: (i) home settings, (ii) school classrooms, and (iii) public places. Evaluation of the robot's performance includes its ability to detect and identify humans in these environments, its ability to express emotions, and its ability to interact with

humans naturally. Different features and capabilities of the robot were used in varying degrees in the three different scenarios and evaluated correspondingly. The article is structured as follows: In Section 2, the hardware and software components of the robot are presented. We emphasize the description of the functionalities that allow the robot to provide human-like communication capabilities and to be empathic. Section 3 describes the robot's applicability and functionality in three different natural environments. Finally, in Sections 4 and 5, a discussion and some conclusions of this work are given.

2 Bender: A General-Purpose Social Robot

The main idea behind the design of Bender (Ruiz-del-Solar et al., 2011) was to have an open and flexible research platform that provides human-like communication capabilities, as well as empathy. Bender has an anthropomorphic upper body (head, arms, chest) and a differential-drive platform that provides mobility (see Figure 1). The electronic and mechanical hardware components of the robot, as well as its software architecture, are described in the following paragraphs.

Figure 1. Bender, a general-purpose social robot. The gripper arm is mounted only when required (e.g., for cleaning tasks).

2.1 Hardware

The main hardware components of the robot are as follows:

- **Chest.** The robot's chest incorporates an HP 2710p tablet PC as the main processing platform. This device is powered by a 1.2-GHz Intel Core 2 Duo processor with 2 GB of DDR2 667 MHz memory, running the Windows XP Tablet PC edition operating system. The tablet includes 802.11bg connectivity. The screen of the tablet PC allows (i) the display of relevant information

Figure 2. Bender's facial expressions.

for the user (a web browser, images, videos, etc.) and (ii) data entry by way of the touch-screen capability.

- **Head.** The robot's head incorporates two cameras, one standard CCD camera (Philips ToUCam III SPC900NC) in the left eye and one thermal camera (FLIR TAU 320 thermal camera (FLIR TAU 320)) in the right eye. The cameras are connected to the tablet PC using USB ports. The head is capable of pan-tilt movements and has several servomotors that move the mouth, eyebrows, and antennae-like ears. In addition, a circular array of red, green, and blue (RGB) LEDs is placed around each eye. The head, mouth, eyebrows, and antennae-like ear movements, as well as the array of RGB LEDs are controlled using dedicated hardware (PIC18F4550-based), which communicates with the tablet PC via a USB port. The head's weight is about 1.6 kg. The robot combines different cues to show emotions. The two eyebrows move independently, and the jaw moves up and down. The two antennae also move independently, and the array of LEDs around the eyes is used with different color configurations to simulate moods. There are three different colors (red, green, and blue), and there are three groups of LEDs—one for each color—that can be switched on and off independently. Figure 2 shows some of the non-verbal expressions that Bender is able to show: happy, surprised, sad, and angry. Note that these expressions are dynamic (e.g., the movement of the antennae; see <http://vision.die.uchile.cl/videos/> for some example videos).

- **3D Vision.** The robot is equipped with a Kinect motion sensing input device placed below the robot's chest (see Figure 1). This position allows the device to detect persons and also objects while grasping.

- **Arms.** The two arms of the robot were designed to allow it to manipulate objects. They are strong enough for raising a large glass of water or a filled coffee cup. Each arm has six degrees of freedom: two in the shoulder, two in the elbow, one in the wrist, and one for the gripper. The actuators consist of eight servomotors (six RX-64 and two RX-28). The arms are controlled directly from the tablet PC via a USB port. Each arm weighs about 1.8 kg.

- **Gripper arm.** A gripper arm was designed to allow the robot to manipulate objects at the floor level (see Figure 1). The gripper arm has five degrees of freedom; two in the shoulder, one in the elbow, one in the wrist, and one for a dual-opening gripper. The actuators consist of five servomotors (three RX-64 for the shoulder and elbow, and two RX-28 for the wrist and the gripper). The gripper arm is also controlled directly from the tablet PC via a USB port. The gripper arm weighs about 0.8 kg. It is important to note that this arm is mounted on the robot only in applications where it is required to manipulate objects at the floor level (e.g., cleaning). In most human-robot interaction applications, such as the ones described in this paper, the gripper arm is not mounted.

- **Mobile platform.** All described structures are mounted on a mobile platform. The platform is a Pioneer 3-AT, which has four wheels, provides skid-steer mobility, and is connected to a Hokuyo URG-04LX laser for sensing. This platform is endowed with a Hitachi H8S microprocessor. A Dell Alienware notebook, powered by an Intel Core i7 processor with 6 GB of DDR3 memory and running Linux, is placed on top of the mobile platform with the task of

running the navigation software. A second Dell Alienware notebook, running Windows 7, is in charge of running the vision and manipulation modules. All computers are connected using a local area network (Ethernet 100 Mbps).

2.2 Software Architecture

LLC - Low-Level Control: Runs in a dedicated hardware.

Figure 3. Software architecture. (Top) Low-level control processes running on dedicated hardware. All high-level processes run on the tablet PC or on the notebooks. (Bottom) The hardware components: chest, head, arm and hand, and platform.

The main components of our software architecture are shown in Figure 3. Speech synthesis and analysis, as well as vision tasks (general object recognition; face, hand, and gesture recognition), are performed on the Windows 7-based Dell Alienware notebook, whereas the Navigation and Mapping modules are run from the other Dell Alienware notebook (running Linux). The Low-Level Control modules run on dedicated hardware (head and arm control). Both notebooks are connected using URBI, an open-source software platform for controlling robots (<http://www.urbiforge.org/>). The main goals of URBI are to help *make robots compatible* and to *simplify the process of writing programs and behaviors*. All the modules running on the HP 2710p tablet PC are controlled through URBI using UObjects (a C++ component library available in the URBI standard API).

The Navigation and Mapping modules that provide self-localization, mapping, path planning, and collision avoidance, among other functionalities, are implemented using the Robot Operating System (ROS, <http://www.ros.org/>), an open-source meta-operating system for robots. ROS provides different services, including hardware abstraction, low-level device control, implementation of commonly used functionalities, message passing, and package management. It

also provides tools and libraries for obtaining, building, writing, and running codes across multiple computer platforms.

The Speech Analysis and Synthesis module provides a speech-based interface to the robot. Speech recognition is based on the use of several grammars suitable for different situations. The Speech Synthesis module uses Festival's text-to-speech tool, dynamically changing certain parameters between words in order to obtain human-like speech. This module is implemented using a control interface built with the CSLU toolkit. The CSLU toolkit (<http://www.cslu.ogi.edu/toolkit>), developed by the Center for Spoken Language Understanding of the Oregon Health & Science University, uses a graphical user interface for constructing simple spoken dialogue systems. Similarly, the Vision module provides a visual interface to the robot. This module is implemented using our own algorithms (Correa, Hermosilla, Verschae, & Ruiz-del-Solar, 2012; Correa, Ruiz-del-Solar, Verschae, Lee-Ferng, & Castillo, 2010; Hermosilla, Ruiz-del-Solar, Verschae, & Correa, 2012; Ruiz-del-Solar, Verschae, & Correa, 2009). The latest addition to this module is the robust detection and recognition of faces/people using a visual camera and a thermal camera (Correa et al., 2012; Hermosilla et al., 2012). It is important to note that the robot can operate without the use of the thermal camera, but at the cost of slightly reducing its human-detection capabilities, which, in particular, could increase the number of false detections generated by the vision module. This is particularly important for scenarios in which many subjects are present at the same time.

The primary task of the Low-Level Control module is to generate control orders to the robot's head, arms, gripper arm, and mobile platform. The Emotions Generator module is in charge of generating the specific orders corresponding to each emotion. Emotions are generated in response to specific situations within the finite-state machine that implements high-level behaviors. Finally, the Strategy module is in charge of selecting the high-level behaviors to be executed. First, the user selects which high-level behavior will be used. This is usually done using the tablet PC on the chest of the robot. Each high-level behavior is previously implemented using a script that indicates the actions and the behavior to be performed by the robot. The definition of the behavior takes into account what sensory inputs will be taken by the robot and what kind of action will be performed, such as speaking, displaying information via the tablet PC, or searching the Internet for information. This script has information about which modules to load, defines the basic functions of the behavior, and, finally, has the sequence of rules that describe the behavior. Most behaviors simply define a sequence of actions to be done after an input from the user (either by speech, a tap on the touch screen of the tablet PC, or after a person or face has been detected by the Vision module). Of special interest for this article are the capabilities for face and hand analysis included in the Vision module. The Face and Hand Analysis module incorporates the following functionalities:

- face detection using boosted classifiers (Verschae & Ruiz-del-Solar, 2003; Verschae, Ruiz-del-Solar, & Correa, 2008),
- face recognition using a histogram of local binary pattern (LBP) features computed on the visual and thermal cameras (Correa et al., 2012; Hermosilla et al., 2012; Ruiz-del-Solar et al., 2009),
- person detection using thermal and visual information, as well as face information (Hermosilla et al., 2012),
- people tracking using face information and Kalman filters (Ruiz-del-Solar, Verschae, Vallejos, & Correa, 2007),
- gender classification using facial information (Verschae, Ruiz-del-Solar, & Correa, 2006), and
- age classification using facial information, hand detection using skin information, and recognition of static hand gestures (Correa et al., 2010; Francke, Ruiz-del-Solar, & Verschae, 2007).

Bender's most important functionalities are listed in Table 1. All these functionalities have already been tested successfully as single modules. Table 2 shows a quantitative evaluation of the

Table 1. Bender's main functionalities.

Capability	How achieved	Source/reference
Mobility	Differential-drive platform (Pioneer 3-AT)	http://www.mobilerobots.com
Speech recognition and synthesis	CSLU toolkit	http://www.cslu.ogi.edu/toolkit
Face detection and recognition	Face and hand analysis module	(Hermosilla et al., 2012; Ruiz-del-Solar et al., 2009; Verschae et al., 2008)
Person detection and recognition	Person detection module	(Ruiz-del-Solar et al., 2010)
Hand gesture recognition	Face and hand analysis module	(Correa et al., 2010)
General purpose object recognition	SIFT-based object recognition module	(Loncomilla & Ruiz-del-Solar, 2008)
Robot detection	Object recognition module	(Arenas, Ruiz-del-Solar, Norambuena, & Cubillos, 2009)
Robot soccer refereeing	Object recognition module	(Arenas et al., 2009)
Expression of emotions	Anthropomorphic mechatronics head (7 DOF)	Own group hardware and control software
Object manipulation	Two arms (6 DOF), one gripper arm (5 DOF)	Own group hardware and control software
Information visualization	The robot's chest sports a 12-inch display	Standard tablet PC
Standard computer inputs	Keyboard and mouse	
Other computer inputs	Touch screen on the chest and virtual keyboard	Standard tablet PC
Internet access	802.11b connectivity	

Note. DOF = Degrees of freedom.

human-robot interaction functionalities, measured using standard databases. As can be observed in these databases, the obtained results are among the best reported. This is an important issue because we would like our social robot to use the best tools and algorithms when interacting with people. For instance, we do not want the robot to have problems when detecting people who are immersed in an environment with variable lighting conditions.

2.3 Comparative Analysis

Several social robots were created in the last decade (Breazeal & Scassellati, 1999; Dongwoon et al., 2011; Metta, Sandini, Vernon, Natale, & Nori, 2008; Gouaillier et al., 2009; Stückler et al., 2012; Kaneko et al., 2009, 2011; Movellan, Tanaka, Fortenberry, & Aisaka, 2005; Rusu, Holzbach, Diankov, Bradski, & Beetz, 2009; Sakagami et al., 2002; Tellez et al., 2008; Chen et al., 2011), all of which focusing on different features and achieving important goals in problems ranging from manipulation of objects to human-like interaction.

In order to interact with humans, two very important features or capabilities that a humanoid robot should possess are non-verbal communication (Tojo, Matsusaka, Ishii, Kobayashi, 2000; Ekman & Friesen, 1969) and capacity to evoke empathy. Robots implementing non-verbal communication capabilities tend to express themselves by moving facial parts (mostly the mouth, eyebrows, and eyelids), as well as their body. To be able to evoke empathy, a humanoid robot not only should be humanlike, but also should give the impression of familiarity, but without being eerie (MacDorman & Ishiguro, 2006); all these fall under the uncanny valley hypothesis (Mori, 1970). Appearance and behavior are the two main factors accounting for familiarity and eeriness.

Table 2. Evaluation of some of bender's functionalities measured using standard databases.

Problem	Database	Results	Comments
Face detection ^a			
– Single face	BioID	DR=95.1%, FP=1	Best results reported
– Single face	FERET	DR=98.7%, FP=0	No report
– Multiple faces	CMU-MIT	DR=89.9%, FP=25	No report
– Multiple faces	UCHFACE	DR=96.5%, FP=3	No report
Face tracking ^b	PETS-ICVS 2003	DR=70.7%, FP=88 (set A) DR=70.2%, FP=750 (set A)	Best results reported
Eyes detection ^a			
– Single face	BioID	DR=97.8%, MEP=3.02	Best results reported
– Single face	FERET	DR=99.7%, MEP =3.69	No report
– Multiple faces	UCHFACE	DR=95.2%, MEP =3.69	No report
Gender classification ^a			
– Single face	BioID	CR=81.5%	No report
– Single face	FERET	CR=85.9%	No report
– Multiple faces	UCHFACE	CR=80.1%	No report
Face recognition			
– Standard test ^c	FERET fafb	Top-1 RR=97%	Among best results reported
– Variable illumination ^d	YaleB	7 individuals per class, Top-1 RR=100% 2 individuals per class, Top-1 RR=96.4%	Best results reported
– Variable illumination ^d	PIE	2 individuals per class, Top-1 RR=99.9	Best results reported
Hand gesture recognition ^e	Group's own database, real-world videos, 4 static gestures	RR=70.4%	No report
Person detection ^f	Group's own database and real-world settings	DR=86.8%, FP=0 (real-world settings)	No report

^aReported in Verschae et al. (2008).^bReported in Ruiz-del-Solar et al. (2007).^cReported in Ruiz-del-Solar et al. (2009).^dReported in Ruiz-del-Solar and Quinteros (2008).^eReported in Francke et al. (2007)^fReported in Ruiz-del-Solar et al. (2010).

Note. DR = Detection rate; FP = number of false positives; RR = recognition rate; CR = classification rate; MEP = mean error in pixels; no report = no other reports in the same dataset.

Some existing robots are not well suited as social robots (e.g., Cosero and PR-2) because they largely lack the ability to evoke emotions (because they do not have facial expressions, and/or their body expression is highly limited) and also because they lack familiarity. Other robots have a human-like body (e.g., Reem-B, KeJia, HRP-4, ASIMO, NAO), but they completely lack facial

expressions. However, some of them can evoke empathy in humans by means of their body expressions (Beck, Hiolle, Mazel, & Cañamero, 2010).

On the one hand, robots with realistic human faces, such as HRP-4C and EveR-3, can reproduce human-like facial expressions, but they might be eerie. This is because they are human replicas that look and act almost, but not exactly, like actual human beings, thus causing a negative impression on humans (Mori, 1970; MacDorman & Ishiguro, 2006).

On the other hand, robots with non-realistic human-based faces, such as iCub, Kismet, and Nexi, that can reproduce facial expressions and evoke empathy in humans are better suited for social robotics from the point of view of evoking empathy.

The Bender robot described in this manuscript is a humanoid robot with manipulation and human-like interaction capabilities. Its simple design and expressive face evoke empathy on the users and invite interaction with it, as will be shown in the experiments reported in the next section. Some of the main differences with existing robots regarding its expressiveness are the use of antennae-like ears, the use of an array of colored LEDs around the eyes to simulate moods, and the up-down movement of its jaw. As with most robots expressing emotions, Bender has eyebrows that can change direction and can move its head up and down and from left to right.

Emotions are complex states of the mind, which include social roles and cognitive factors. Some emotions are considered basic (Ekman, 1992), as they are not reducible to others. Basic emotions seem to be fundamentally universal, and their external manifestation seems to be independent of culture and personal experience (Lane & Nadel, 2000). For this reason, they can be revealed by facial expressions without the intervention of verbal language (Ekman & Friesen, 1978). Six basic emotions, with corresponding facial expressions, have been identified: anger, surprise, disgust, sadness, happiness, and fear. In our robot, four of these basic emotions are implemented: anger, surprise, sadness, and happiness.

Moreover, colors play a fundamental role in giving the viewer information about the nature of objects. Colors can also stimulate the viewer to respond in a particular way. Da Pos and Green-Armytage (2012) found a relationship between emotions, facial expressions, and colors in studies using subjects from Australia and Europe; their conclusions therefore may not apply to other cultures. For the four basic emotions implemented in Bender, the corresponding results observed by Da Pos and Green-Armytage are as follows:

- red is the main color related to anger,
- blue is the main color related to sadness,
- yellow, green, red, and blue are the main colors related to surprise, and
- yellow and red are the main colors related to happiness.

There are similarities between the results of Da Pos and Green-Armytage (2012) and the colors used by Bender for three emotions: anger, sad, and surprised. The colors used in happiness are different; Bender uses only green, whereas Da Pos and Green-Armytage show that yellow and red are the colors related to happiness. Happiness is the most difficult emotion to achieve in Bender owing to its incapability to smile.

3 Applicability in Natural Environments

3.1 Home settings

One of the main goals of the development of our social robot is its use as an assistant and companion for humans in their home environments. The idea is that the robot should be able to interact freely with non-expert users in these places. Of course, we realized that we needed to conduct a lengthy process before we could achieve this goal. In 2006, we decided that a very appropriate way to achieve this was to participate regularly in the RoboCup@Home competition. RoboCup@Home focuses on real-world applications and human-machine interaction with autonomous robots in home settings. Tests are related to manipulation of typical objects that can be found in a home-like environment, with navigation and localization inside a home scenario and

with interaction with humans. Our social robot has participated in the RoboCup@Home world competition since 2007, and in 2007 and 2008, it received the RoboCup@Home Innovation Award as the most innovative robot in the competition. The technical committee members of the league decided this award. The robot's capabilities that were appreciated the most were its empathy, ability to express emotions, and human-like communication capabilities.

Bender has also had successful participation in the "Who's Who?" benchmark defined in the RoboCup@Home competition. The main goal of this benchmark is to test the ability of domestic robots "to detect and recognize people autonomously in an unknown environment." In order to accomplish this task, it is expected that "without manual calibration, a robot will introduce itself to a group of people, ask for their names, memorize them, and recognize the people when meeting them again." The test focuses "on human detection/recognition, face detection/recognition, safe navigation, and human-robot interaction with unknown people."

Basically, the test is as follows: The robot enters the arena through the door and stops next to it. Two people enter through the door and introduce themselves to the robot, one by one. The robot asks for their names and memorizes them. During this first interaction, the robot has to speak to the subject, detect the subject's face, and store the face information to recognize him/her later. When told to do so by an operator, the robot goes to the room and starts looking for guests. In the room, there are two other persons who are unknown to the robot. One of them is sitting and the other one is standing. There is also one person standing in the room who is known to the robot. When the robot finds a person, it has to approach and interact with that person in order to get his attention and perform facial recognition, and then indicate whether it has identified that person. In case of positive identification, the robot must state the person's name; otherwise, it states that the person is unknown. The distance from the robot to the person must not exceed 1 m. The robot works autonomously through the whole test.

Correa et al. (2012) presented the quantitative results of the use of Bender for solving the "Who's Who?" benchmark, using the functionalities described in Table 2 in a realistic domestic scenario (see Figure 4 for an example). On average, Bender was able to detect 86.6% of the unknown and known persons present at the scene and to determine the identity of the known persons in 92% of the cases.

3.2 Classroom Setting

Robotics is a highly motivating activity for children. It allows them to approach technology in both an amusing and an intuitive manner, while discovering the underlying scientific principles. Indeed, robotics has emerged as a useful tool in education because, unlike many other areas, it provides a venue where scientific and technological fields or ideas intersect and overlap (Ruiz-del-Solar & Aviles, 2004). With the objective of studying the use of social robots as a tool for fostering the interest of children in science and technology, we decided to employ our social robot as a lecturer for school children. Since 2007, the robot has given several talks to school children of different ages in different classrooms or auditoriums in different locations.

We shall describe a concrete experience in which the robot gave talks to school children who were 10–13 years old. Altogether, 228 school children participated in this activity, and each time one complete class attended the talk in a multimedia classroom. (More than 10 talks were given by the robot.) The duration of each talk was 55 min, and it was divided into two parts. In the first part, the robot introduced itself and talked about its experiences as a social robot. In the second part, the robot explained some basic concepts about renewable energies and about responsible energy use. After the talk, the students were able to interact freely with the robot, thanks to its face detection and speech recognition capabilities. During these interactions, the robot behaved autonomously, as it did while giving the talk. The talk was given using the multimedia capabilities

Figure 4. Environment of a domestic setting where Bender was evaluated (Correa et al., 2012) following the “Who’s Who?” experimental setup of the RoboCup@Home competition. Examples of (a) a visible image, (b) a thermal image, and (c) an output image showing face detections (blue), face candidate detections (green), and people detections (red). Views from the robot and from above the stage are shown in (d) and (e), respectively.

Table 3. Children's evaluation of the robot's presentation.

Question topic	Excellent (%)	Good (%)	Regular (%)	Bad (%)	Very bad (%)
Robot presentation	59.6	28.1	11.4	0.9	0.0

of the robot: speech generation and multimedia presentation (a Microsoft PowerPoint-like presentation). The presentation can be displayed on the robot's chest using the tablet PC or using a standard video projector. In the experiments presented in the following paragraphs, the presentation was given using a standard video projector. During the talk, the robot used mainly its speech generation capabilities and it moved its body to address and face all the audience.

After the robot's lecture, the children, without any previous notification, answered questions from a survey regarding their personal appreciation of the robot as well as regarding some specific topics discussed by the robot. In the robot evaluation part, the children were asked to give an overall evaluation of the robot. On a linear scale of grades going from 1 to 7 (7 being best), the robot was given an average score of 6.4, which is about 90%. In the second part, the children evaluated the robot's presentation by choosing one of five options: Excellent, good, regular, bad, and very bad. As shown in Table 3, more than half of them chose excellent. The third question was, "Do you think it is a good idea for robots to teach some specific topics to schoolchildren in the future?" For which 92% of the children answered *yes*. In the technical content evaluation part, five questions were asked using the multiple choice test method with five choices in each case. The first three questions were related to energy sources (classification of different energy sources as renewable or non-renewable, availability of renewable sources, and indirect pollution produced by renewable sources). The fourth question asked about the differences between rechargeable and non-rechargeable batteries, and the fifth question asked about the benefits of efficient energy use. The percentage of correct answers to these five questions is shown in Table 4. The average percentage of correct answers was 55.4%. All these percentages are better than random choice (20%).

Table 4. Results of the technical content evaluation.

Question topic	Energy source classification	Renewable sources	Pollution	Batteries	Benefits of energy use	Avg.
Correct answers, %	75.9	33.7	31.6	75.0	60.6	55.4
Confidence interval (95%)	$\pm 5.54\%$	$\pm 6.10\%$	$\pm 6.00\%$	$\pm 5.60\%$	± 6.36	± 6.40

In summary, we observed that children gave a very good evaluation of the robot (6.4 out of 7) and that 87.7% of them evaluated the presentation as excellent or good. They also had very favorable opinions about the use of robots as lecturers in a classroom environment (92%). Moreover, the children were able to learn some basic technical concepts (the overall percentage of correct answers was 55.4%), although they had just heard them once from a robot. The main goal of this evaluation of the technical content was just to see whether the children could learn some basic information from the robot and not to measure how well they learned it. Therefore, control experiments with human instructors were not carried out at that time. This will be part of a future work. Finally, it is important to note that the robot was able to give its talk and to interact with the children without any human assistance.

During the last 5 years, Bender has given talks to more than 2,000 school children. The talks have been given in classrooms and laboratories to groups of 20–25 children and in a big auditorium to more than 200 children in one session (see Figure 5).

Figure 5. Bender giving talks to school children in an auditorium and in different classrooms and laboratories.

3.3 Public Setting

We tested the applicability of our social robot in a public setting (Ruiz-del-Solar et al., 2010). The main idea of the experiment was to let humans interact freely with the robot, using only speech and visual cues (face and hand gestures, facial expressions, etc.). The robot was placed in some public spaces inside our university campus (mainly halls of buildings), and the students passing

Figure 6. Bender interacting with students in a public space inside the university.

through these public spaces could interact freely with the robot if they wanted to (see Figure 6). The robot did not move by itself during the whole experience in order to avoid any risk of collision with the students; therefore it needed to catch people's attention just by using speech, visual cues, and other strategies such as complaining about being alone, being bored, or calling out to people. When the robot detected a student in its vicinity (by using the face and person detection modules), it asked the student to approach it and to have a little conversation: The robot presented itself, asked some basic information about the student, and, afterwards, it asked the student to evaluate its capabilities in expressing emotions. Finally, after the evaluation, the robot thanked the student and the interaction was finished. In the whole process, the robot always behaved autonomously, and the development team did not intervene or communicate with the robot.

During the experiment, the ability of the robot to express emotions was evaluated using the following procedure: The robot expressed an emotion randomly, and then it asked the student to identify it. The student gave its answer using the touch screen (choosing one of four alternatives). This process was repeated four times to allow the student to evaluate the four different emotions. This means that if the user had selected an expression randomly, a 25% recognition rate would have been achieved.

To take advantage of the multimedia capabilities of the robot, we asked users to give their answers using the touch screen to ensure that speech recognition mistakes would not affect the experiment. This was the only instance in which the interaction between the robot and a human was not based on speech or visual cues. No external human assistance was given to the robot's users at any time, and the robot always behaved autonomously; the development team observed the situation from several meters away. After the human-robot interaction had finished, and the humans had left the robot's area, they were asked to evaluate their experience with it using a poll.

Our first observation was that, from the total number of students who passed near the robot, about 37% curiously approached it. Thirty-one percent of those who approached the robot interacted with it; the rest, however, just observed it. The total number of students who interacted with the robot was 83. The age range was 18 to 25 years old, and the gender distribution was 70% males and 30% females. Out of the 83 students, 74.7% finished the interaction and 26.3% left before finishing. The main reasons for leaving prematurely were as follows: (i) the students were not able to interact with the robot satisfactorily (because of speech recognition problems; see the

Table 5. Recognition rate of the robot's facial expressions.

Expression	Correctness	Confidence interval (95%)
Happy	51.0%	$\pm 13.23\%$
Angry	76.5%	$\pm 11.52\%$
Sad	78.4%	$\pm 11.23\%$
Surprised	76.5%	$\pm 11.52\%$
Average	70.6%	

Discussion section), (ii) they did not have enough time to complete the emotion evaluation, or (iii) they were not interested in making the evaluation. The mean interaction time with the subjects who finished the interaction, including the emotion evaluation, was 124 s.

In Table 5, the recognition rate of the different expressions is shown. It can be observed that the average expression recognition rate was 70.6%, and that all expressions, except “happy,” have a recognition rate greater than 75%. In Tables 6 and 7, the results of the poll, taken by the users after interacting with the robot, are presented. It should be emphasized that only 74.7% of the users who finished the interaction with the robot answered the poll. As can be seen in Tables 6 and 7, 83.9% of the users judged the robot's appearance to be excellent or good; 88.5% evaluated the robot's ability to express emotions as excellent or good; and 80.7% rated the robot's ability to interact with humans as excellent or good. In addition, 90% of them thought that it was easy to interact with the robot; 84% believed that the robot would be suitable to be a receptionist, museum guide, or butler; and 67% thought that the robot could be used for educational purposes with children. It should be mentioned that the whole experiment was carried out inside an engineering campus, and, therefore, the participants were engineering students with high probability of enjoying technology and robots. On the other hand, we believe that, as expert users in technology, they can be more critical about robots than standard users.

Table 6. Human evaluation of the robot's appearance and interaction abilities.

	Excellent	Good	Regular	Bad	Very bad
Robot appearance	30.7%	53.2%	14.5%	1.6%	0%
Ability to express emotions	31.1%	57.4%	8.2%	3.3%	0%
Ability to interact with humans	17.8%	62.9%	17.7%	1.6%	0%

Human facial expression recognition by both humans and computers has been widely studied in the past (for recent studies, see Bettadapura, 2012; Fasel & Luetten, 2003; Pantic & Rothkrantz, 2000). It is generally agreed that there are six “archetypal” expressions that are innate and universal (Ekman, 1994; Izard 1994). They correspond to happiness, sadness, anger, surprise, disgust, and fear. When recognized using computer vision systems, it has been shown that the

Table 7. Human evaluation of the robot's applicability and ease of use.

	Yes	No
Do you think that it is easy to interact with the robot?	90%	10%
Do you think that the robot is suitable to be a receptionist, museum guide or butler?	84%	16%
Do you think that the robot can be useful in tasks related to interactions with children?	67%	33%

expressions of surprised, happiness, and disgust have very high recognition rates (100%, 93.4%, and 97.3%, respectively) because of the clear motion of the mouth and the eyebrows (Pardas & Bonafonte, 2002). In the case when humans perform the recognition, it has been shown (Sebe et al., 2007) that anger is usually confused with disgust and that fear is usually confused with surprise. This is attributed to the fact that there are common facial movements and actions being shared by these expressions. It is important to remember that the four facial expressions implemented by Bender are part of the six archetypal expressions that are innate and universal (Ekman, 1994; Izard, 1994). The two expressions not implemented by Bender are disgusted and scared. In Bender, the mouth has only one degree of freedom (for moving the jaw up and down), and, as mentioned, in humans the movement of the mouth is the more important cue in expressing/recognizing the expressions of surprise and happiness. The implemented surprise expression in Bender is based on an opening of the mouth (by moving the jaw), which gives a good enough response. On the other hand, movement of the jaw (and the inability to smile) does not clearly imply happiness, and the use of green LEDs around the eyes seems to imply more a neutral/calm expression than a happy expression.

4 Discussion

4.1 Evaluation Methodology

There are different approaches to evaluating the performance of social robots when interacting with humans. Although the performance of isolated algorithms should be measured (e.g., recognition rate of a face recognition algorithm), it is also necessary to analyze the effect robots have on humans. Some researchers have proposed employing quantitative measures of the characteristics of human attention [attitude (Reeves & Nass, 1996), eye gaze (Ono, Imai & Ishiguro, 2001), etc.] or of the body movement interactions between a human and a robot (Kanda et al., 2004). We firmly believe that acceptance and empathy are two of the most important factors to be measured in a human–robot interaction context and that these factors can be measured using survey-based methods that express the users’ opinions. The social robot described has been evaluated by about 300 people with different backgrounds (228 school children, 62 engineering students, and 5 international researchers from the RoboCup@Home competitions), which indicates the validity of the results obtained and the conclusions reached.

4.2 Evaluation of the Robot’s Capabilities

As can be observed in Table 2, the vision-based human–robot interaction functionalities of the robot measured using standard databases are reported to be among the best. This is very important because the robot must have robust tools and algorithms to deal with dynamic conditions in the environment. In addition, the robot has received two innovation awards from the scientific community specializing in domestic service robots, which indicates that the robot theoretically is able to interact with people adequately.

4.3 Robot Evaluation When Interacting With People

In our experiments with children in a real classroom setting, we learned that the children gave very good evaluations of the robot, with 87.7% of them evaluating its presentation as excellent or good. They also have a very favorable opinion about the use of robots as lecturers in a classroom environment. We can conclude that the robot was well accepted by this group of children (10–13 years old) who had the opportunity to interact with a robot for the first time. Bender was able to give its talk and to interact with the children without any human assistance. In addition, the children were able to learn some basic technical concepts from it (55.4% correct answers to five technical questions, where five alternative choices were given for answering each question).

It should be stressed that the robot’s presentation was a standard lecture, without any repetition of its contents. In addition, it should be noted that the robot, unlike a human teacher, cannot detect distracted children in order to call their attention nor can it achieve the same level of expressivity either in speech or in gestures, leaving it only with its empathy and other

mechanisms, such as moving the mouth while talking, to catch the listener's attention. These results encourage us to explore further the relevance of an appealing human-robot interaction interface. Naturally, it will be necessary to carry out a comparative study of the performance of robot teachers against human teachers and to analyze the dependence of the results on the specific topics to be taught (technical topics, foreign languages, history, etc.).

In our experiments in public settings, we tested the ability of the social robot to interact with people freely. The experiments were conducted in the halls of different buildings inside our engineering school. Thirty-seven percent of the students passing near the robot approached it, and 31% of them interacted directly with it. In all cases, the robot tried to attract the students actively by talking to them. It was interesting to note that 26.3% of the students who interacted with the robot left before finishing the interaction. One of the main reasons for leaving was that the students were not able to interact comfortably with the robot, because of its speech recognition problems. Our speech recognition module has limited capabilities; it is not able to recognize natural language, and the recognition is perturbed by environmental noise. This is one of the main technical limitations of our robot and, in general, of other service robots as well. Nevertheless, 74.7% of the students completed the emotion evaluation that the robot presented to them, with a mean interaction time of 124 s.

Before carrying out these experiments, we had the intuitive and qualitative impression that the emotions our robot could generate were adequate and that a human could understand them. The quantitative evaluation obtained in the experiments showed us that this perception was correct and that humans can correctly recognize the robot's expressions in 70.6% of the cases. We are in the process of designing a new "happy" expression (for work in progress, see http://vision.die.uchile.cl/videos/very_happy.mp4) in order to replace the evaluated happy expression that was recognized in only 51% of the cases. In our opinion, this expression is still too close to the "neutral expression," basically because of the shape and movement of the mouth, which has only one degree of freedom. The array of LEDs around the eyes and the antennae have balanced this lack of movement of the mouth and seem to be rather effective for the other emotions. Nevertheless, we will replace the current LED array with a new one that can generate a larger amount of color arrangements. In addition, we expect to improve the emotion expression capabilities of the robot. Although the mechanics of the robot's head impose some limits on the expressions that it can generate (limited by the number of degrees of freedom in the face and by the shape and movement of the jaw), we believe the current expressions are rich enough to produce empathy in the users. This confidence has been verified by all the reported experiments and also by unreported interactions between the robot and external visitors to our laboratory.

Acceptance of the robot by the engineering students, as it was in the case of the children, was high (83.9% evaluated the robot's appearance as excellent or good; 88.5% rated the robot's ability to express emotions as excellent or good; and 80.7% judged the robot's ability to interact with humans to be excellent or good). In addition, 90% of the students thought that it was easy to interact with the robot, and 84% and 67% of the students thought that the robot could be used as a household assistant or for educational purposes, respectively.

Currently, Bender very frequently participates in public technology trade fairs and events for promoting technology among children and the general public (see Figure 7).

As no single feature or capability of the robot can solely account for an effective and affective communication with humans, we plan to improve some of Bender's features that seem to affect this communication; this is part of our future research. In particular, some of the modules that need further work and analysis are the speech recognition module, the speech generation module, and the emotion generation module of the happy expression. In addition, there are several features and modes of interaction that need further investigation and analysis, such as gaze, expression, and recognition of gender and age.

Figure 7. (a–b) Bender participating in trade fairs and events. (c–d) INAPI Trade Fair (Santiago, Chile, July 2009), Salamanca City Exposition (Chile, Nov. 2011)

5 Conclusions

The main goal of this article was to report and analyze the applicability of a general-purpose social robot in three different natural environments: (i) home settings, (ii) school classrooms, and (iii) public spaces. Evaluation of the robot's performance depends on the robot's ability to detect and identify humans in those places, its ability to express emotions, and its ability to interact naturally with humans using human-like expressions. The experiments show that the robot is able to interact successfully with humans using human-like interaction mechanisms, such as speech and visual cues (especially face information). It is notable that children learned some facts from the robot despite its limitations. Nevertheless, this aspect should be further analyzed.

From a technical point of view, the visual-based human–robot interaction functionalities of the robot, as measured using standard databases, are among the best, and the robot has received two innovation awards from the scientific community, which indicate that Bender is able to interact with people adequately. However, one of its main technical limitations is the speech recognition module, which needs to be improved.

As future work, we would like to analyze further the teaching abilities of our robot. In general terms, we believe that more complex methods should be used to measure how much the children learn from the robot and how this learning compares with that from a human teacher.

Acknowledgments

This research was partially funded by FONDECYT Projects 1090250 and 3120218 (CONICYT, Chile).

References

- Arenas, M., Ruiz-del-Solar, J., Norambuena, S., & Cubillos, S. (2009). A robot referee for robot soccer. *Lecture Notes in Computer Science* 5399 (RoboCup Symposium 2008) 426-438. http://dx.doi.org/10.1007/978-3-642-02921-9_37
- Beck, A., Hiolle, A., Mazel, A., & Cañamero, L. (2010, October). Interpretation of emotional body language displayed by robots. In *Proceedings of the 3rd international workshop on Affective interaction in natural environments*, 37-42. ACM. <http://dx.doi.org/10.1145/1877826.1877837>
- Bettadapura, V. (2012). Face Expression Recognition and Analysis: The State of the Art, *Tech Report*, *arXiv:1203.6722*
- Breazeal, C., & Scassellati, B. (1999), How to build robots that make friends and influence people, in: *Proceedings of the International Conference on Intelligent Robots and Systems*, 2, 858-863. <http://dx.doi.org/10.1109/IROS.1999.812787>.
- Correa, M., Hermosilla, G., Verschae, R., & Ruiz-del-Solar, J. (2012). Human Detection and Identification by Robots using Thermal and Visual Information in Domestic Environments, *Journal of Intelligent and Robotic Systems*, 66 (1-2), 223-243. <http://dx.doi.org/10.1007/s10846-011-9612-2>
- Correa, M., Ruiz-del-Solar, J., Verschae, R., Lee-Ferng, J., & Castillo, N. (2010). Real-Time Hand Gesture Recognition for Human Robot Interaction. *Lecture Notes in Computer Science* 5949 (RoboCup Symposium 2009), 46-57. <http://dx.doi.org/10.1109/LARS.2009.5418324>
- Chen, X., Jin, G., Ji, J., Wang, F., & Xie, J. (2011), KeJia Project: Towards Integrated Intelligence for Service Robots, *Technical Report*, USTC-CS-MAS-2011-H1, Multi-Agent Systems Lab., Department of Computer Science of China, HeFei, China.
- Choi, D., Lee, D.W., Lee, D.Y., Ahn, H.S., & Lee, H. (2011). Design of an android robot head for stage performances. *Artif. Life Robot.* 16, 3 (December 2011), 315-317. <http://dx.doi.org/10.1007/s10015-011-0938-4>
- Da Pos, O., & Green-Armytage, P. (2012). Facial expressions, colours and basic emotions. *JAIC- Journal of the International Colour Association*, 1.
- Ekman, P. & Friesen, W. V. (1978). *Manual for facial action coding system* (Palo Alto: Consulting Psychologists Press, 1978).
- Ekman, P. (1992). *Cognition and Emotion*, 6, 169–200.
- Ekman, P. (1994). Strong evidence for universals in facial expressions: A reply to Russell's mistaken critique. *Psychological Bulletin*, 115, 268-287. <http://dx.doi.org/10.1037//0033-2909.115.2.268>
- Ekman, P., & Friesen, W.V.(1969), The repertoire of non- verbal behavior, *Semiotica*, 1, 49-98.
- Fasel, B., & Luetttin, J. (2003), Automatic facial expression analysis: A survey, *Pattern Recognition*, 36, 259–275. [http://dx.doi.org/10.1016/S0031-3203\(02\)00052-3](http://dx.doi.org/10.1016/S0031-3203(02)00052-3)
- FLIR TAU 320 thermal camera. (n.d.) Information available on October 2012 in <http://www.flir.com/cvs/cores/view/?id=51374>
- Franke, H., Ruiz-del-Solar, J., & Verschae, R. (2007). Real-time Hand Gesture Detection and Recognition using Boosted Classifiers and Active Learning. *Lecture Notes in Computer Science* 4872 (PSIVT 2007), 533-547. http://dx.doi.org/10.1007/978-3-540-77129-6_47
- Gouaillier, D., Hugel, V., Blazevic, P., Kilner, Ch., Monceaux, J., Lafourcade, P., Marnier, B., Serre, J., & Maisonnier, B. (2009). Mechatronic design of NAO humanoid, *ICRA '09. IEEE International Conference on Robotics and Automation*, 769-774, (12-17 May 2009). <http://dx.doi.org/10.1109/ROBOT.2009.5152516>

- Hayashi, K., Sakamoto, D., Kanda, T., Shiomi, M., Koizumi, S., Ishiguro, H., Ogasawara, T., & Hagita, N. (2007) Humanoid Robots as a Passive-Social Medium – A Field Experiment at a Train Station –, Proc. Conf. Human-Robot Interaction – HRI’07, Virginia, 137-144, (March 8-11). <http://dx.doi.org/10.1145/1228716.1228735>
- Hermosilla, G., Ruiz-del-Solar, J., Verschae, R., & Correa, M. (2012). A Comparative Study of Thermal Face Recognition Methods in Unconstrained Environments, *Pattern Recognition*, 45, (7), 2445-2459. <http://dx.doi.org/10.1016/j.patcog.2012.01.001>
- Ishiguro, H., Ono, T., Imai, M., & Kanda, T. (2003) Development of an interactive humanoid robot Robovie—An interdisciplinary approach, in Robotics Research, R. A. Jarvis and A. Zelinsky, Eds. New York: Springer-Verlag, 179–191. http://dx.doi.org/10.1007/3-540-36460-9_12
- Izard, C.E. (1994), Innate and universal facial expressions: Evidence from developmental and cross-cultural research, *Psychological Bulletin*, 115 (2), 288-299, (Mar 1994). <http://dx.doi.org/10.1037/0033-2909.115.2.288>
- Kanda, T., Ishiguro, H., Imai, M., & Ono, T. (2004) Development and Evaluation of Interactive Humanoid Robots, Proc. IEEE, 92 (11), 1939-1850. <http://dx.doi.org/10.1109/JPROC.2004.835359>
- Kanda, T., Ishiguro, H., Ono, T., Imai, M., & Nakatsu, R. (2002), Development and evaluation of an interactive humanoid robot Robovie, in Proc. IEEE Int. Conf. Robotics and Automation, 1848–1855. <http://dx.doi.org/10.1109/ROBOT.2002.1014810>
- Kaneko, K., Kanehiro, F., Morisawa, M., Akachi, K., Miyamori, G., Hayashi, A., & Kanehira, N. (2011), Humanoid robot HRP-4 - Humanoid robotics platform with lightweight and slim body, *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 4400-4407, (25-30 Sept.). <http://dx.doi.org/10.1109/IROS.2011.6094465>
- Kaneko, K., Kanehiro, F., Morisawa, M., Miura, K., Nakaoka, S., & Kajita, S. (2009), Cybernetic human HRP-4C, *9th IEEE-RAS International Conference on Humanoid Robots*, 2009..7-14, (7-10 Dec. 2009). <http://dx.doi.org/10.1109/ICHR.2009.5379537>
- Lane, R D. & Nadel, L. (2000). Cognitive neuroscience of emotion (Oxford: Oxford University Press).
- Loncomilla, P., & Ruiz-del-Solar, J. (2008). Robust Object Recognition using Wide Baseline Matching for RoboCup Applications. *Lecture Notes in Computer Science 5001 (RoboCup 2007)* 441-448. http://dx.doi.org/10.1007/978-3-540-68847-1_46
- Macdorman, F.K., & Ishiguro, H. (2006), The uncanny advantage of using androids in cognitive and social science research, *Interaction Studies*, 7 (3), 297-337, <http://dx.doi.org/10.1075/is.7.3.03mac>
- Matsusaka, Y. Tojo, T., Kubota, S., Furukawa, K., Tamiya, D., Fujie, S., & Kobayashi, T. (1999) Multi-person conversation robot using multimodal interface, in *Proc. World Multiconf. Systems, Cybernetics and Informatics*, 7, 450–455.
- Metta, G., Sandini, G., Vernon, D., Natale, L., & Nori, F. (2008) The iCub humanoid robot: an open platform for research in embodied cognition. In Proceedings of the 8th Workshop on Performance Metrics for Intelligent Systems (PerMIS '08). ACM, New York, NY, USA, 50-56. <http://dx.doi.org/10.1145/1774674.1774683>
- Mori, M. (1970). Bukimi no tani (The uncanny valley, Original in Japanese) (K. F. MacDorman & T. Minato, Trans.). *Energy*, 7(4), 33–35
- Movellan, J.R., Tanaka, F., Fortenberry, B., & Aisaka, K. (2005), The RUBI/QRIO Project: Origins, Principles, and First Steps, *Proceedings of The 4th International Conference on Development and Learning*, 80-86, (19-21 July 2005). <http://dx.doi.org/10.1109/DEVLRN.2005.1490948>
- Nakadai, K., Hidai, K., Mizoguchi, H., Okuno, H.G., & Kitano, H. (2001) Real-time auditory and visual multiple-object tracking for robots, in Proc. Int. Joint Conf. Artificial Intelligence., 1425–1432.

- Ono, T., Imai, M., & Ishiguro, H. (2001), A model of embodied communications with gestures between humans and robots, in Proc. 23rd Annu. Meeting Cognitive Science Soc., 732–737.
- Pantic, M., & Rothkrantz, L.J.M. (2000) Automatic analysis of facial expressions: the state of the art, *IEEE Trans. on Pattern Analysis and Machine Intelligence* 22 (12), 1424–1445.
- Pardas, M., & Bonafonte, A. (2002), Facial animation parameters extraction and expression recognition using Hidden Markov Models, *Signal Processing: Image Communication*, 17, 675–688. [http://dx.doi.org/10.1016/S0923-5965\(02\)00078-4](http://dx.doi.org/10.1016/S0923-5965(02)00078-4)
- Reeves, B., & Nass, C. (1996) The Media Equation. Stanford, CA: CSLI.
- ROS website (n.d.): Web site: <http://www.ros.org/wiki>
- Ruiz-del-Solar, J. (2010). Robotics-Centered Outreach Activities: An Integrated Approach. *IEEE Trans. on Education*, 53 (1), 38-45. <http://dx.doi.org/10.1109/TE.2009.2022946>
- Ruiz-del-Solar, J. Verschae, R., Vallejos, P., & Correa, M. (2007) Face Analysis for Human Computer Interaction Applications, Proc. 2nd Int. Conf. on Computer Vision Theory and Appl. – VISAPP 2007, Special Sessions, 23 – 30, Barcelona, Spain.
- Ruiz-del-Solar, J., & Aviles, R. (2004). Robotics Courses for Children as a Motivation Tool: The Chilean Experience. *IEEE Trans. on Education*, 47 (4), 474-480. <http://dx.doi.org/10.1109/TE.2004.825063>
- Ruiz-del-Solar, J., & Quinteros, J. (2008) Illumination Compensation and Normalization in Eigenspace-based Face Recognition: A comparative study of different pre-processing approaches. *Pattern Recognition Letters*, 29 (14), 1966-1979. <http://dx.doi.org/10.1016/j.patrec.2008.06.015>
- Ruiz-del-Solar, J., Correa, M., Smith, F., Hevia-Koch, P., Leottau, L., Pairo, W., & Verschae, R. (2011). UChile HomeBreakers 2011 Team Description Paper, *RoboCup Symposium*, (5-11 July 2011), Istanbul, Turkey (CD Proceedings).
- Ruiz-del-Solar, J., Mascaró, M., Correa, M., Bernuy, F., Riquelme, R., & Verschae, R., (2010). Analyzing the Human-Robot Interaction Abilities of a General-Purpose Social Robot in Different Naturalistic Environments. *Lecture Notes in Computer Science 5949 (RoboCup Symposium 2009)*, 308–319. http://dx.doi.org/10.1007/978-3-642-11876-0_27
- Ruiz-del-Solar, J., Verschae, R., & Correa, M. (2009). Recognition of Faces in Unconstrained Environments: A Comparative Study. *EURASIP Journal on Advances in Signal Processing (Recent Advances in Biometric Systems: A Signal Processing Perspective)*, 2009, 1, <http://dx.doi.org/10.1155/2009/184617>
- Rusu, R.B., Holzbach, A., Diankov, R., Bradski, G., & Beetz, M. (2009), Perception for mobile manipulation and grasping using active stereo, *9th IEEE-RAS International Conference on Humanoid Robots*, 632-638, (7-10 Dec. 2009). <http://dx.doi.org/10.1109/ICHR.2009.5379597>
- Sakagami, Y., Watanabe, R., Aoyama, C., Matsunaga, S., Higaki, N., & Fujimura, K. (2002), The intelligent ASIMO: system overview and integration, *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 3, 2478-2483. <http://dx.doi.org/10.1109/IRDS.2002.1041641>
- Sebe, N., Lew, M.S., Sun, Y., Cohen, I., Gevers, T., & Huang, T.S. (2007), Authentic Facial Expression Analysis, *Image and Vision Computing*, 25, 1856-1863. <http://dx.doi.org/10.1109/AFGR.2004.1301585>
- Stückler, J., Dröschel, D., Gräve, K., Holz, D., Schreiber, M., & Behnke, S. (2012,) NimbRo @Home 2012 Team Description. In *RoboCup 2012 @Home League team descriptions*, Mexico City, June 2012.
- Tellez, R., Ferro, F., Garcia, S., Gomez, E., Jorge, E., Mora, D., Pinyol, D., Oliver, J., Torres, O., Velazquez, J., & Faconti, D. (2008). Reem-B: An autonomous lightweight human-size humanoid robot, *8th IEEE-RAS International Conference on Humanoid Robots*. 462-468, (1-3 Dec. 2008). <http://dx.doi.org/10.1109/ICHR.2008.4755995>
- Tojo, T., Matsusaka, Y., Ishii, T., Kobayashi, T. (2000), A conversational robot utilizing facial and body expressions, *IEEE International Conference on Systems, Man, and Cybernetics*, 2, 858-863. <http://dx.doi.org/10.1109/ICSMC.2000.885957>

- Verschae, R., & Ruiz-del-Solar, J. (2003) A Hybrid Face Detector based on an Asymmetrical Adaboost Cascade Detector and a Wavelet-Bayesian-Detector, *Lecture Notes in Computer Science* 2686, 742-749. http://dx.doi.org/10.1007/3-540-44868-3_94
- Verschae, R., Ruiz-del-Solar, J., & Correa, M. (2006) Gender Classification of Faces using Adaboost, *Lecture Notes in Computer Science* 4225, 68–78. http://dx.doi.org/10.1007/11892755_7
- Verschae, R., Ruiz-del-Solar, J., & Correa, M. (2008) A Unified Learning Framework for object Detection and Classification using Nested Cascades of Boosted Classifiers, *Machine Vision and Applications*, 19 (2), 85-103. <http://dx.doi.org/10.1007/s00138-007-0084-0>
-

Authors' names and contact information: J. Ruiz-del-Solar, Department of Electrical Engineering and Advanced Mining Technology Center, Universidad de Chile, Chile. Email: jruizd@ing.uchile.cl; M. Correa, Advanced Mining Technology Center, Universidad de Chile, Chile. Email: macorrea@ing.uchile.cl; R. Verschae, Advanced Mining Technology Center, Universidad de Chile, Chile. Email: rodrigo@verschae.org; F. Bernuy, Department of Electrical Engineering and Advanced Mining Technology Center, Universidad de Chile, Chile. Email: fbernuy@ing.uchile.cl; P. Loncomilla, Advanced Mining Technology Center, Universidad de Chile, Chile. Email: ploncomi@ing.uchile.cl; M. Mascaró, Department of Electrical Engineering, Universidad de Chile, Chile. Email: mmascaro@ing.uchile.cl; R. Riquelme, Department of Electrical Engineering, Universidad de Chile, Chile. Email: romrique@ing.uchile.cl; F. Smith, Department of Electrical Engineering, Universidad de Chile, Chile. Email: smith1813@gmail.com.