

DEVELOPMENT OF A MINIATURE PAIR OF EYES WITH CAMERA FOR THE HUMANOID ROBOT ROBOTA

Dimitris Pongas, Florent Guenter, Andre Guignard, Aude G. Billard
Autonomous Systems Laboratory (ASL3)
School of Engineering
EPFL
1015 Lausanne, Switzerland
<http://asl.epfl.ch>; <http://robota.epfl.ch>
aude.billard@epfl.ch florent.guenter@epfl.ch

January 6, 2005

Abstract

A growing trend in humanoid robotics tend at reducing the size of humanoids in order to lower their building costs. While growing small has its advantages, it also has drawbacks. In particular, providing miniature humanoids with the same sensorimotor capabilities as their grown-up peers is a challenge, both in terms of mechanics, electronics and control. The Robota project creates humanoids, whose size should match that of a commercial doll, so that they can be used as part of experiments with autistic children. Because these experiments measure the child's ability to socially interact with others, it is fundamental to provide the robot with sensory capabilities, such as speech and vision, that are at the basis of these interactions. This paper presents the creation of a miniature pair of mobile eyes, mounted with miniature cameras to provide Robota with binocular and mobile vision. The system allows the robot to blink, to direct its gaze toward or away from its user and to track the user's face.

1 Introduction

Humanoids seem to grow smaller by the day! Indeed, the most recent high-quality platforms are miniature, see, e.g. Honda Asimo [1], Sony QRIO [2] and Fujitsu HOAP.

There are numerous advantages to go small. Small is cute. Small is safer. Small is cheaper. The latter is probably the most important aspect of the miniaturization of humanoids, as it opens the doors for global commercialization of these robots. However, there are also a number of drawbacks. Miniature components (motors, gears, etc) are not as readily available as larger ones. Miniature controllers remain limited in their processing power. Small implies also light. This limits the choice of batteries, as well as the number of sensor components.

Nevertheless, the miniaturization of humanoids is progressing rapidly. It is certainly advantaged by the fast development of handheld devices, such as

PDA's, cell phones and watches. Telecommunication companies have open a large market for cheap and powerful miniature processors and batteries, sophisticated softwares, such as speech recognition and synthesis, as well as to a range of miniature sensors, such as cameras, wireless/radio/bluetooth transceivers. PDA's have become a practical and cheap means of human-robot interaction, e.g. [3, 4, 5], and, one can take for granted that the number of applications of PDA's and their accessories in robotics will only increase in the future. The Robota project follows such a trend to provide its doll-shaped humanoid robots with state-of-the-art portable processors and sensors [6, 7, 3]. In this paper, we present the recent creation of a binocular tracking systems for Robota, composed of a pair of mobile eyes mounted with cameras.

1.1 The Robota Project

The Robota project builds a series of multiple degrees of freedom doll-shaped humanoid robots, whose physical features resemble those of a small child. The Robota project is part of a current trend of robotics research that develops educational and interactive toys for children with disabilities; see, e.g. [8]. The Robota project started in 1997 with the first prototype. Numerous iterations thereafter followed, leading to the creation of a commercial prototype, sold by DIDEL SA.

Robota's Hardware: The original Robota robot, sold by DIDEL SA, has 5 degrees of freedom (DOF): 1 DOF for each leg, arm and head. A prototype of a 7 DOFs miniature arm has been constructed to extend Robota's original capabilities. For a complete description of mechanics of these components, see [9].

Robota's electronic components consist of a Motor Board and a Sensor Board. The Motor board is addressed via a RS232 serial interface from a PC or a PocketPC. The Sensor Board is addressed from the Motor Board via a SPI (serial peripheral interface) serial interface. Motor Board and Sensor Board are controlled by two PIC16FA microcontrollers. Schematic of the electronic boards and of the mechanics are available in the Robota UserGuide at <http://robota.epfl.ch>

Robota's behaviors allow multi-modal human-robot interactions, including the ability to imitate gesture, to dialog via speech processing/synthesis, and to learn simple composition of motions and of words, see [6, 10].

Applications: The Robota robots find an application in two areas: 1) in education, as a programming tool at the University, and as an entertaining interactive toy in museums [6]; and 2) in behavioral studies with autistic children, that investigate the potential of an imitator robot to test the ability of autistic children to imitate and to teach these children some simple coordination behaviors [11, 12]. These experiments aim at evaluating autistic children's ability to socially interact with others. In particular, they contrast the effect of the human-like features of Robota on the willingness of the child to interact in comparison to the effect of a more robot-like set of features. Hence, in order to make the interaction with Robota comparable to that a child could have with a human, it is fundamental to provide Robota with sensory capabilities, such as speech and vision, that are at the basis of these interactions. The development of the new prototype of mobile eyes, described in this paper, aims at fulfilling this need. With a mobile pair of eyes, Robota will be able to direct or avert its



Figure 1: A PocketPC mounted with a FlyCam Camera tracks the motion of the user and sends commands to the microcontrollers and servos of Robota, so that it can imitate the motion of the head and arms of a user.

gaze from the user. Children with autism usually avert their gaze from others; it will be interesting to determine whether the human-like features affect this behavior or not. Robota will be able to blink, a natural means of transmitting jokes. It will be interesting to see whether autistic children are or can become sensitive to this means of communication in the robot. Finally, thanks to the cameras mounted in the eyes, Robota will be able to track the user's face; a feature it already has through the PocketPC camera, see [3]. With the cameras mounted in the eyes, the tracking will be more precise and will increase the angle of view of Robota.

2 Design of the camera

Miniaturization is a key constraint in that particular project. Robota eyes measure 16mm of diameter (for the smallest version of Robota) and 26mm for the largest ones. These are commercially available glass-made doll eyes. The goal is to be able to insert the camera within the eye. The volume of the eyes is, however, too small to contain the electronic board (that proceeds to the digital conversion of the image). Thus, the sensor must be connected to the board through a *flex* cable in such a way that the cable is not impeding the movement of the eye. The cable must, therefore, be small. Hence, the number of line used for the transmission must be minimal. Our first task was, therefore, to find a chip that would transmit the camera information with as few lines as possible and with a protocol that could be handled by the microcontroller and the PocketPC driving Robota (I2C, SPI, serial or USB).

2.1 A typical camera

A typical color camera sends its image in the RGB or the YUV (otherwise known as YCbCr) format. The RGB format sends 3 components (green, red and blue) for each pixel. With these 3 layers the entire color range can be produced.

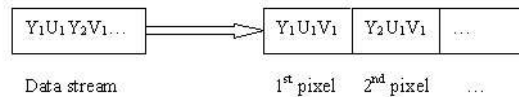


Figure 2: Conversion between YUYV and YUV format

The YUV is a format that was first developed for color televisions in order to be compatible with the black and white. It is widely used throughout Europe. Y stands for luminance (or luma), U(Cb) corresponds to the color difference on the blue and V(Cr) is the color difference for red. In black and white televisions only the Y component is shown. There are more formats but these two are the most used.

In both formats, the response of one pixel has to be encoded in 3 bytes. One way to reduce the number of bytes required for the transmission is to sample the component, i.e. send each pixel component one at a time. The serial transmission for YUV is 4:2:2, and for RGB 6:5:5, ...

Another type of the YUV 4:2:2 format is the YUYV format, where V corresponds to the luminance. The conversion to the normal YUV format is simple: the U and V components are constant for two pixels. An example is presented in figure 2. Using such a compression means a loss of color information, but this is nearly imperceptible by the human eye. In this mode, the transmission for every pixel uses only two bytes.

A typical micro camera uses 8 bits for parallel transmission of the raw image and a couple of bits for the protocol. Each image is composed, usually, of 100K to 1.25M pixels. A VGA (640x480, 300K pixels) output requires 600 KB per image in the YUYV format. For 15 images per second, this results in 9 MB of data per second. For two cameras, this equates to 18 MB per second. Normally a USB 1.1 can handle up to 1.25 MB per second. This means that, before the images can be sent to a computer, a more powerful compression has to be done (JPEG, wavelet, ... for images and MPEG2, Divx, ... for video).

2.2 Data transfer

Even when using a compression factor of 20, the result is around 1MB per second. The rate remains too high to be handle by one of the classical protocols used in robots, I²C, SPI, serial or parallel communication. Only USB communication could handle such a transfer rate. This transfer mode of the images was chosen for the project. The camera must, therefore, be directly connected to the PC or PocketPC.

2.3 List of possible cameras

Thanks to the recent and wide use of cameras for cell phones, it was possible to find a number of cameras compatible with the size of Robota's eyes (see tables 1,2,3, 4). These cameras have, also, the advantage to be provided with the optic part integrated. Moreover, most of these are already provided with a flex cable.

Unfortunately, the mobile phone market is a booming, but quite unpredictable market. Hence, when trying to obtain more information about these

	MB86S02A	LZ0P390	HAM490002	TCM8230MD
Manufacturer	FUJITSU	SHARP	HITACHI	TOSHIBA
Module size(mm)	7.8x6.98x3.95	7x6.5x6	7x7.6x4.9	6x6x4.5
Optical size(in)	1/7	1/7	-	1/6
Image size	CIF, QCIF	CIF, QCIF QVGA QQVGA SubQCIF	-	VGA, QVGA QQVGA CIF, QCIF SubQCIF
Frame rate(fps)	30	30	15	30
Output	Digital	Digital	Digital	Digital
Color/b&w	Color	Color	Color	Color
Color filter	Included	Included	Included	Included
Controlled	I ² C	-	-	I ² C
Power cons.(mW)	30	45	-	40

Table 1: Possible cameras (1/4)

	VS6552	VS6502V/F	CSC1011	CSV1061
Manufacturer	STMicroelectronics		ROHM	
Module size(mm)	5.8x10.56x8.66	10.6x8.7x5.8	8.4x12.7x4.8	8.2x8.2x5.9
Optical size(in)	-	-	1/7	1/4
Image size	VGA	VGA, QVGA	CIF	VGA
Frame rate(fps)	30	60	30	30
Output	Digital	Digital	Digital	Digital
Color/b&w	Color	Color	-	-
Color filter	Included	-	-	-
Controlled	I ² C	I ² C	-	-
Power cons.(mW)	50	20mA	32	68

Table 2: Possible cameras (2/4)

	ITM-M-SC	ITM-C-SL	MCX18LG0A	DS-310 E-B-G
Manufacturer	-	-	SONY	EURECA
Module size(mm)	16x16x15	22x22x28	18.3x18.3x8.4	30x15x11
Optical size(in)	1/4	1/3	-	1/7
Image Size	EIA, CCIR	PAL, NTSC	-	CIF
Frame rate(fps)	-	-	-	-
Output	Analogic	Analogic	Digital	USB
Color/b&w	b&w	Color	Color	-
Color filter	Not Included	Not Included	Included	-
Controlled	-	-	-	-
Power cons.(mW)	50	50	-	-

Table 3: Possible cameras (3/4)

cameras, we encountered numerous problems and unexpected answers! For instance, we were told that the MB86S02A camera from Fujitsu was no longer supported in Europe, that the LZ0P390 needed a minimum order quantity of

	TC5470	TC5747	OV764X
Manufacturer	TRANSCHIP		OMNIVISION
Module size(mm)	9.8x9.45	9.8x9.45	10x10
Optical size(in)	1/4	1/4.5	1/4
Image Size	VGA, QVGA, SubQVGA	VGA, QVGA	VGA, QVGA
Frame rate(fps)	1-20, 1-40	1-20, 1-40	30,60
Output	Digital(JPEG)	Digital(JPEG)	Digital
Color/b&w	Color	Color	Color
Color filter	Included	Included	Included
Controlled	I ² C	I ² C	OV519
Power cons.(mW)	85	85	-

Table 4: Possible cameras (4/4)

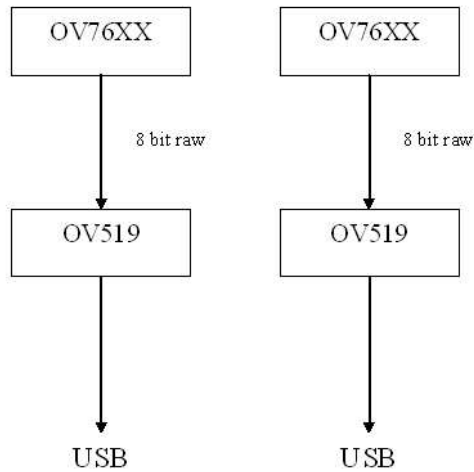


Figure 3: Solution using two OV764X cameras and two chips (OV519) in order to have USB output

500 and had a shipping time of 20 weeks, that the minimum quantity for the TCM8230MD was of 100K units per month (!). We received no response for the HAM490002, VS6552 and VS6502V/F, and the CSC1011 and CSV1061 are still under development. Finally, it turned out that the cameras ITM-M-SC, ITM-C-SL, MCX18LG0A and DS-310 E-B-G were not appropriate (too big) for the application.

The good news, finally, came from the TC574X and the OV764X. The TC574X could be obtained in small quantities (\$30-35). These cameras have also the advantage to output in a compressed(JPEG) format. The OV764X cameras could be purchased in small quantities (10) for 40€. This camera, when coupled with the OV519 chip, has a USB output and it compresses the image using a JPEG compression.

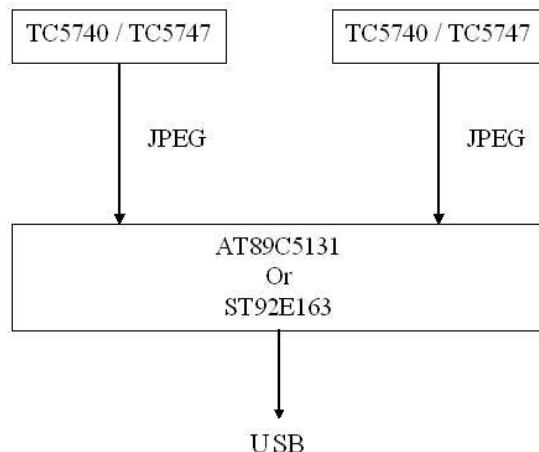


Figure 4: Solution using two TC574X cameras and a chip in order to have an USB output

2.4 Solutions

We, therefore, came out with two possible solutions to our problem. The first solution consists of an OV764X camera sensor coupled with the OV519 chip. The OV519 gets the data uncompressed from the sensor, then compresses them with a JPEG compression and sends them over the USB. For this solution a reference design was obtained. A similar solution is used for the EyeToy camera for the PlayStation. The drivers for the chip exist under both Linux and Windows.

The second solution consists of two TC547X camera that are coupled to the AT89C5131 chip from Atmel and the ST92E163 from STMicroelectronics. Both of these have a USB2 output and the SDK C++ routines that send data over the USB are available. However, the drivers required for retrieving and converting the images do not exist and need to be written.

2.5 Comparison of the solutions

We went finally for the first solution because of the easy implementation and the fact that the drivers existed. Note that solution 2 would still be preferable, since it would have required less chips, hence a smaller PCB. However, everything would need to be programmed: the interface between the chip and the camera, the interface between the chip and the PC (Driver), a module for decompressing and also a module to make the video stream compatible with the viewers have to be written (or adapted).

3 Getting the images

The OV519 and the OV7648FB are communicating through SCSB. It is a communication protocol created by Omnicision and strongly inspired by I²C. The SCSB lines are SIO-C and SIO-D, and go from the OV519 to the connector

	Solution 1	Solution 2
Image sensor	OV764X	TC574X
Chip	OV519	AT89C5131 or ST92E163
Number of outputs	2	1
Drivers	Exist	Do not exist
Implementation difficulty	Easy	Difficult
Price	100€	80€
Reference design	Yes	No
Interface sensor/chip	Ready	ToDo
Number of chips	2	1

Table 5: Comparison table of the two solutions for the camera module.

for the camera. The line SCSB is needed when more than one slave is connected on the bus. The images are transmitted through the 8 parallel bits Y1-Y8. The remaining bits are configuration bits and power lines. The module is powered via USB, hence no external supply is needed. The camera and the chip 3.3V and 2.5V respectively. These are provided by two voltage regulators. The schematic of the module for this configuration can be found in [13].

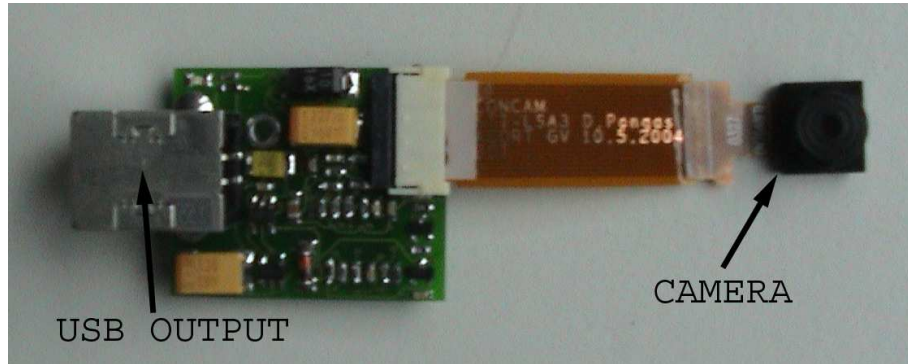


Figure 5: Camera module showed from above

The OV519 offers the possibility to take snapshots. This option was not implemented as it was not needed for our particular application. In addition, the chip can handle a microphone. This particularity will be exploited in further development of the module.

Because of Robota general application in schools, the exploitation system of the robot has to be windows (for PocketPC). There exist two drivers for Windows. The first one is a confidential driver from Omnivision! The second one is the same version of this driver modified for another webcam that uses the OV519!! This driver can be downloaded from the Internet. To be compatible with the configuration of this project some modifications needed to be done. For more details, please, refer to [13].

Once these modifications had been done, the camera worked perfectly. The modified windows driver worked and the quality of image was similar to that of an image taken from a regular webcam, see Figure 7. Also the driver has an

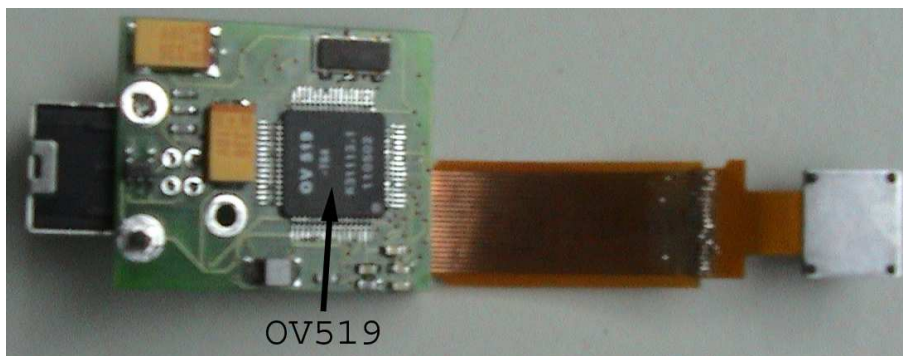


Figure 6: Camera module showed from below



Figure 7: Photo of the 7DOF prototype of Robota arm taken by Robota eye camera

output on the format VFW¹. This format is very standard and can be used by nearly all the programs.

4 Mechanical schema

A prototype of three DOF Eyes was constructed. One DOF drives the horizontal rotation of the two eyes and the two other DOFs drive the vertical rotation of each eye. For a complete description of the mechanical design of the eye, see [9]. The eyes can move vertically and horizontally by about 80 degrees in each direction. The difficult aspect of the mechanical design was to place the center of rotation at the center of each eye, in order for the eyes to move easily within the original doll head.

The principal constraint in this project is the aesthetic of the robot. The only visible part of the prototype is the eyes (the rest of the mechanics being hidden inside the doll's head). We use doll eyes that we modify to insert the

¹Video For Windows

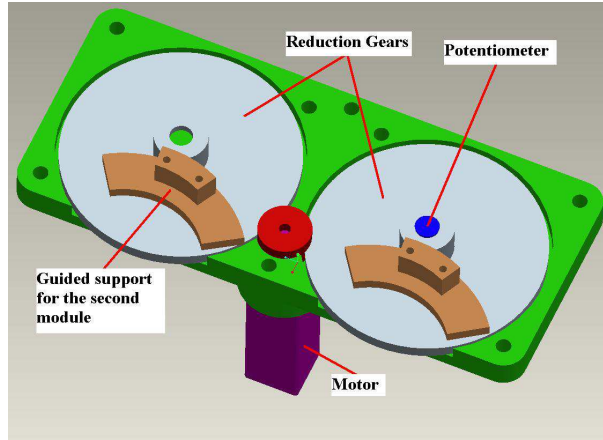


Figure 8: Horizontal rotation of the eyes

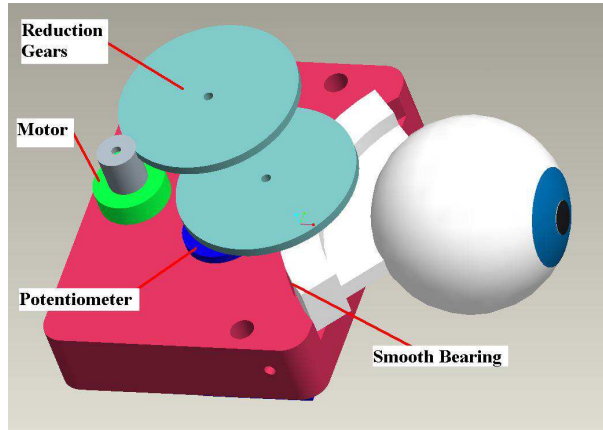


Figure 9: Vertical rotation of the eyes

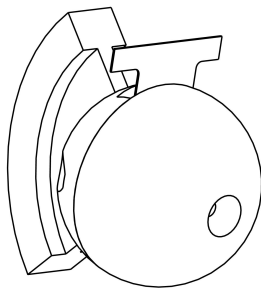


Figure 10: An eye of Robota with the camera

cameras, by drilling a tiny hole through the pupil, making sure that the iris remains intact.

The only way to put the cameras (figure 10) inside the eye is by separating the eye into two parts and put the camera between them (figure 11). The negative of the camera must fill in the space in the back of the iris of the first half of the eye. A tiny hole inside the pupil allows enough light to enter the camera lens without obstructing its angle of view. The second half is fixed to the mechanism of movement. A hole in its upper part allows the cable of the camera to get out. The two parts are fixed using tiny screws.

The electronic circuit of the camera must be placed as close as possible to the eye, and attached to the axis of the first DOF of the eyes. With this construction, the PCB moves when the eyes are making an horizontal movement but stays immobile when they are moving vertically. A *Flex*, i.e. very flexible, cable must join the camera and the PCB in order not to disturb the vertical movement.

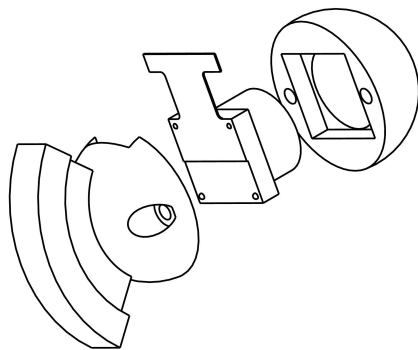


Figure 11: The 3 parts of an eye of Robota with a camera inside

5 Discussion and Conclusion

This paper presented a prototype of a miniature 3 degrees of freedom pair of mobile eyes mounted with cameras. To our knowledge, this is the sole example of mobile visual system of this size and for which all the electronics and mechanics

are contained inside the prototype.

While the majority of large size humanoid robots have binocular eyes, see, e.g., ARMAR at the University of Karlsruhe [14], WABIAN at Waseda university [15], Saika at Tohoku university [16] or MIT's Cog, all small size humanoid robots have at most 1 "eye", i.e. one camera mounted directly in the head, the head having at most 2 DOFs. None of the state-of-the-art miniature humanoid robots have decoupled eyes. Note that a prototype of 3 DOFs articulated neck for Robota is under construction.

6 Acknowledgments

This work was supported in part by the Swiss National Science Foundation, through grant no 620-066127 of the SNF Professorships program and by the European Commission Division FP6-IST Future and Emerging Technologies under Contract FP6-002020, EU Integrated Project COGNIRON.

References

- [1] Y. Sakagami, R. Watanabe, C. Aoyama, S. Matsunaga, and K. Higaki, N. Fujimura, "The intelligent asimo: System overview and integration," in *IEEE/RSJ Intl Conference on Intelligent Robots and Systems*, 2002.
- [2] M. Fujita, Y. Kuroki, T. Ishida, and T. Doi, "A small humanoid robot sdr-4x for entertainment applications," in *IEEE/ASME Intl Conference on Advanced Intelligent Mechatronics*, 2003.
- [3] C. S. and B. A., "Pda interface for humanoid robots," in *Third IEEE International Conference on Humanoid Robots (Humanoids 2003)*, Munich and Karlsruhe, Germany, October 2003.
- [4] T. Fong, C. Thorpe, and B. Glass, "Pdadriver: A handheld system for remote driving," in *IEEE International Conference on Advanced Robotics*, 2003.
- [5] C. Lundberg', C. Barck-Holst, J. Folkesson, and H. Christensen, "Pda interface for a field robot," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, vol. 3, 2003.
- [6] A. Billard, "Robota: Clever toy and educational tool," *Robotics & Autonomous Systems*, vol. 42, pp. 259-269, 2003.
- [7] —, "Drama, a connectionist architecture for on-line learning and control of autonomous robots: Experiments on learning of a synthetic proto-language with a doll robot," *Industrial Robot*, vol. 26:1, pp. 59-66, 1999.
- [8] C. Plaisant, A. Druin, C. Lathan, K. Dakhane, K. Edwards, J. M. Vice, and J. Montemayor, "Storytelling robot for pediatric rehabilitation," in *ASSETS'00*, 2000.
- [9] F. Guenter, A. Guignard, L. Piccardi, and A. Calzascia, M. Billard, "Development of a miniature articulated arm and pair of eyes for the humanoid robot robota," in *IEEE/APS Mechatronics and Robotics 2004*, 2004.

- [10] A. Billard and G. Hayes, "Drama, a connectionist architecture for control and learning in autonomous robots," *Adaptive Behavior, Vol. 7:1*, pp. 35–64, 1999.
- [11] B. Robins, K. Dautenhahn, R. te Boekhorst, and A. Billard, "Effects of repeated exposure of a humanoid robot on children with autism," In *S. Keates, J. Clarkson, P. Langdon and P. Robinson (eds), Designing a More Inclusive World. London: Springer-Verlag*, pp. 225–236, 2004.
- [12] K. Dautenhahn and A. Billard, "Games children with autism can play with robota, a humanoid robotics doll," in *In Proceedings of the 1st Cambridge Workshop on Universal Access and Assistive Technology, March*, 2002.
- [13] D. Pongas, "Mobile cameras for the eyes of robota," *Report of senester project, EPFL/STI/I2S/ASL 3*, 2004.
- [14] K. Berns, T. Asfour, and R. Dillman, "Armar-an anthropomorphic arm for humanoid service robot," in *IEEE Intl Conference on Robotics and Automation*, 1999.
- [15] H. Lim and A. Takamishi, "Waseda biped humanoid robot realizing human-like motion," in *IEEE Intl Conference on Systems, Man, and Cibernetics*, 2000.
- [16] A. Konno, N. Kato, S. Shirata, and M. Furuta, T. Uchiyama, "Development of a light-weight biped humanoid robot," in *IEEE/RSJ Intl Conference on Intelligent Robots and Systems*, 2000.