Physical Interaction Learning: Behavior Adaptation in Cooperative Human-Robot Tasks Involving Physical Contact

Shuhei Ikemoto*
JST, ERATO and
Adaptive Machine Systems
Osaka University
Osaka, Japan

Heni Ben Amor
VR and Multimedia Group
TU Bergakademie Freiberg
Freiberg, Germany

Takashi Minato
JST, ERATO
Asada Project
Osaka, Japan

Hiroshi Ishiguro
JST, ERATO and
Adaptive Machine System
Osaka University
Osaka, Japan

Bernhard Jung
VR and Multimedia Group
TU Bergakademie Freiberg
Freiberg, Germany

***Abstract*— In order for humans and robots to engage in direct physical interaction several requirements have to be met. Among others, robots need to be able to adapt their behavior in order to facilitate the interaction with a human partner. This can be achieved using machine learning techniques. However, most machine learning scenarios to-date do not address the question of how learning can be achieved for tightly coupled, physical touch interactions between the learning agent and a human partner. This paper presents an example for such human in-the-loop learning scenarios and proposes a computationally cheap learning algorithm for this purpose. The efficiency of this method is evaluated in an experiment, where human care givers help an android robot to stand up.**

I. Introduction

Robot technology has come a far way from large, unsafe manufacturing machines, to highly sophisticated androids with human-like appearance. As this technology continues to improve, the application domains of robots also keep coming closer to our everyday life. So far, the most common type of robots, namely industrial robots have primarily inhabited dedicated working environments in factories. For a human, entering such a workspace can result in severe injuries. Recent robotic developments, however, are more and more targeted at domestic environments and assistive tasks, where human-robot interaction is indispensable. For humans and robots to share a common living environment, several requirements need to be met. First, all physical contact between the interaction partners needs to be safe, in particular meaning that the human being is never harmed. Next, the robot needs to be able to adapt its behavior to the environment and the actions of the human partner. Ideally, the robot should also learn from previous interaction experiences and modify the behavior according to received critiques.

Learning and adaptation has been intensively studied in the robotics community. In particular, imitation learning has proved to be a promising way of teaching new skills without the need for tedious manual programming. However, research in this area mostly considers scenarios where teacher and learner rely on communication interfaces such as motion capture or use simple symbolic communication. In contrast, recent research such as [5] has considered the direct physical interaction between the communication partners, based on kinesthetic and haptic feedback. While these works constitute an important step towards the meaningful coexistence and interaction of humans and robots, the robot is mostly assumed to take a passive role during the learning task.

In this paper we introduce a physical human robot interaction scenario with a tight coupling between the human instructor and the learning robot. Inspired by the parenting behavior observed in humans, a test subject is asked to physically assist a state-of-the-art robot in a standing up motion. Both human and robot need to adapt their behavior, such that they can cooperatively solve the task. In particular, this also means that the robot needs to react appropriately to the force applied by the human instructor. After each trial, the human can judge whether the interaction was successful or not and the resulting critique is used by a machine learning algorithm to update the behavior of the robot. As learning progresses, the robot creates a behavioral model, which implicitly includes the actions of the human counterpart. To ensure that safety is always guaranteed, the robot is equipped with pneumatically actuated flexible-joints. The robot joints have a high flexibility in response to externally applied force and allow for both passive and active reactions.

We argue, that human-in-the-loop learning scenarios, such as the one presented here, will be particularly interesting in the future, as they can help to strengthen the mutual relationship between humans and robots. Ideally, this will lead to a higher acceptance of robotic agents in our society.

II. Related Work

Important aspects of Physical Human Robot Interaction (PHRI) have been investigated in a perspective research project conducted by European Network of Excellence (EURON) [3]. The project's objective was to lay out and dis-

*Corresponding author. Email: ikemoto@ed.ams.eng.osaka-u.ac.jp

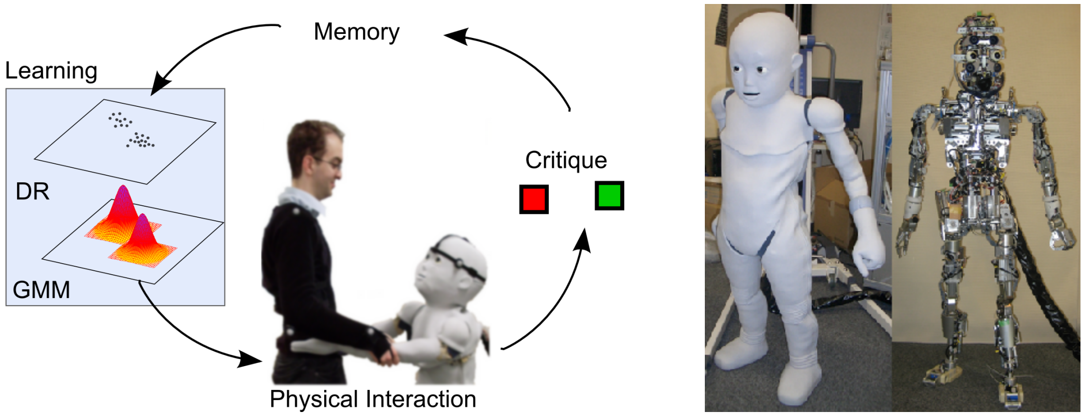

Fig. 1. Left: Overview of the physical interaction learning approach used in this paper: after physical interaction the human judges whether the interaction was successful or not. This information is stored in the memory and later used for learning. Right: the flexible-joint humanoid robot used in the experiments of this paper.

cuss important requirements for safe and dependable robots involved in PHRI. Initial approaches to achieving these requirements are currently being addressed in a follow-up research project called PHRIENDS[1] . The primary goal of the project is to design robots that are intrinsically safe, in order to reduce risks and fatalities in industrial manufacturing workplaces. For this, new actuator concepts, safety measures and control algorithms are being developed, which take the presence of human subjects into account. The results of this project are also relevant to applications outside the manufacturing industry. However, learning and adaptation between humans and robots is not in the focus of the project.

In [7], Khatib et al. discuss basic capabilities needed to enable robots to operate in human-populated environments. In particular, they discuss how mobile robots can calculate collision free paths and manipulate surrounding objects. For this, they characterize free space using an elastic strip approach. The described robots, however, were not expected to get into direct (physical) contact with surrounding human subjects. The importance of direct physical interaction was highlighted in the Haptic Creature project [10]. The project investigates the role of affective touch in fostering the companionship between humans and robots. Another attempt to close human-robot interaction is the work presented in [6]. Here, Kosuge et al. present a robot that can dance with a human by adaptively changing the dance steps according to the force/moment applied on it. In [1], Berger et al. use kinesthetic interactions to teach new behaviors to a small humanoid robot. Additionally, the behavior can be further optimized with respect to a given criterion in simulation. In this learning scheme the robot is a purely passive interaction partner and only acts after learning is finished. Similar approaches to teaching new skills have also been employed in [2] and [9] using different learning methods, i.e. Continuous Time Recurrent Neural Networks and Gaussian Mixture Models respectively. Odashima and colleagues developed a robot that can have direct physical contact with humans. The robot is intended for care tasks, such as carrying injured persons to nearby physicians. The robot can also learn new behaviors and assistive tasks by observing human experts performing them. However, this learning does not take place during interaction, but in offline sessions using immersive virtual environments. In this paper we present experiments with a soft skin robot that is involved in close physical interaction with a human caregiver. In contrast to the above research, both human and robot play an active role in the interaction. Further, they both learn to adapt their behavior to the interaction partner so as to achieve a common goal. This tight coupling of robot and human learning and co-adaptation is a distinctive feature and is the main focus of the work to be presented.

III. Physical Interaction Learning Approach

The goal of interaction learning is to improve the cooperation of humans and robots *while* they are working to achieve a common goal. In Figure 1 we see an overview of the learning scheme employed in this paper. After an initial physical interaction between a human and a robot, the human is given the chance to evaluate the behavior of the robot. More precisely, the human can judge whether the interaction was successful or not. The feedback can be done in various ways, such as through touching or through a simple graphical user interface. Once the critique information is collected by the robot system, it is stored in a memory database. The memory's task is to collect information about recent successful interactions and manage the data for the later learning step. The idea is based on the human short-term memory. It allows us to optimize the set of training examples used for learning, in order to improve learning quality.

After a number of interactions, the learning system queries the memory for a set of new training data. The data is then projected onto a low-dimensional manifold using dimensional reduction techniques. There are three justifications for this step. First, dimensional reduction allows it for a reduction of the space in which learning takes place. From this follows that learning can be much faster and

[1] Physical Human-Robot Interaction which is Dependable and Safe

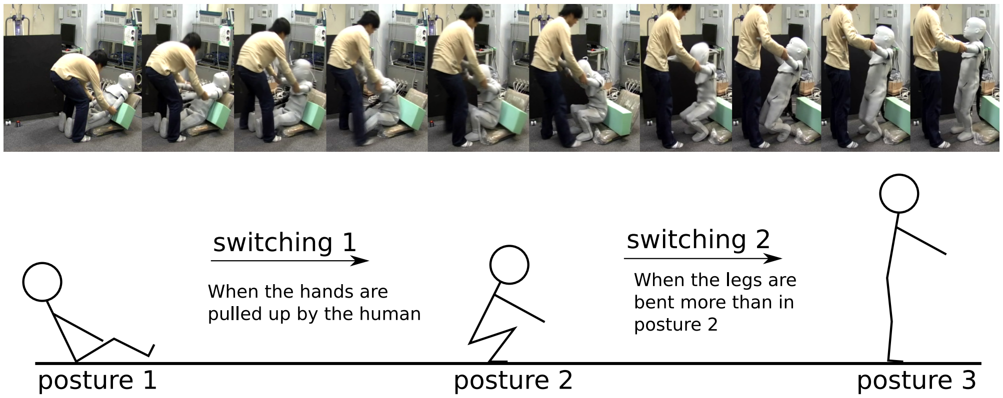

Fig. 2. The three desired postures used in the stand-up behavior of the experiment. The task of learning is to determine ideal switching conditions between the desired postures.

more efficient. Although the system described here does not aim at biological plausibility, it can still be argued that dimensional reduction also takes place in the human brain. Biologically inspired neural networks such as Self-Organizing Maps perform dimensional reduction and create topographic maps such as those found in the human and animal brain. Dimensional reduction generally helps to detect meaningful low-dimensional structures in high-dimensional inputs. Second, dimensional reduction allows us to visualize and understand the adaptation taking place during interaction. This is particularly helpful for later review and analysis purposes. Finally, it also reduces the negative influence of outliers on learning. The inputs to the dimensional reduction step are high-dimensional state vectors describing the robot's postures during the interaction. Output is a low-dimensional posture space.

Once the state vectors are projected onto a low-dimensional manifold, we group the resulting points into sets according to the action performed in that state. Thus, we get for each possible action a set of states in which the corresponding action should be triggered. For each action, a Gaussian Mixture Model is learned. The model encodes a probability density function of the learned state vectors. By computing the likelihood of a given state vector p in a GMM of action A, we can estimate how likely it is that the robot should perform action A when in posture p. The learned models are then used during the next physical interaction trial to determine the robot's actions. For this, each new posture is projected into the low-dimensional posture space. Then, the likelihood of the projected point for each GMM is computed. Following a maximum-likelihood rationale the action corresponding to the GMM with the highest likelihood is then executed by the robot.

With each iteration of the above learning loop, the robot's adapts it's model more and more towards successful interactions. The result is a smoother and easier cooperative behavior between the human and the robot.

A. The CB^2 Robot

The robot used in this study is called Child-robot with Biomimetic Body or CB^2[8]. The robot has the following features:

- It is 130 cm high and weighs about 33 kg.
- It has 56 degrees of freedom (DOFs).
- All joints, apart from the joints used to move the eyes and eyelids, are driven by pneumatic actuators.
- All joints, apart from the joints used to move the fingers, have potentiometers.

The joints have low mechanical impedance due to the compressibility of air. The joints can also be made completely passive if the system discontinues the air compression during robot motion. This helps the robot to perform passive motion during physical interaction and helps to ensure the human helper's safety. This is in contrast to most other robots, where the joints are driven by electric motors with decelerators. Due to the flexible actuators, the joints produce smooth-looking motion, even when the input signal changes drastically. This feature of the CB^2 robot is used to realize complex motions using a simple control architecture. More specifically, full body motions of the robot are realized by switching between a set of successive desired postures. Each posture is described by a posture vector $\boldsymbol{x}$, with each entry of the vector denoting the angular value of a particular joint. A low-level controller is implemented by PD-control of angular values. Each time the desired posture is drastically switched, large drive torques are generated, resulting in active force applied to the human caregiver. As the robot's posture approaches the desired posture, the passive motion gradually becomes the dominant motion of the robot.

In Figure 2 we see how the rising-up behavior used throughout this paper is realized in our control architecture. The behavior is realized by switching between three desired postures. At first glance this approach renders the specification of the robot motion extremely simple. However, the switching times are highly dependent on the human interaction counterpart. More specifically, the switching times depend on the anatomy and skills of the human. This means that the robot has to adapt the switching times to his partner's characteristics while the interaction is going on.

B. Learning Method

In the uprising task explained in this paper, the goal of learning is to determine an ideal timing for switching actions between different desired postures. This is achieved by learning three different probabilistic low-dimensional posture models, one for each desired posture $s \in S$.

At each time step of an interaction between a human and the robot, the posture of the latter and the current desired posture is recorded. The robot posture r is a 52 dimensional vector coding the current angular value of each joint. After the interaction is finished, the postures are stored in the memory. The memory database holds information of the last 10 interactions. Although there are many possible ways how new data is integrated in to the database, the general policy used here is:"new data overwrites old data, successful interactions overwrite failed interactions".

After 10 interactions, the training data from the memory is used for learning. First, dimensional reduction is applied on the data. While many methods can be applied for this task, we used PCA in this paper. To perform PCA, the mean r_m is subtracted from all recorded posture vectors and the covariance matrix M of the resulting points is computed. A singular value decomposition (SVD) on M yields matrices U,V and W, such that:

$$M = UWV^T \tag{1}$$

The columns of matrix V contain orthonormal vectors called the *eigenvectors* or *principal components* of matrix M. The matrix W is a diagonal matrix containing the singular values. Each principal component (PC) has a corresponding singular value which indicates how much information of the data set it covers. The first few PC's are then used as the axes of our lower dimensional PCA space. Given a new data point we can compute its coordinates in PCA space by subtracting the mean and calculating the dot product for each of the principal components.

Next, we compute a GMM for each of the three desired postures. For this we divide the projected data points into distinct sets, according to which desired posture s each point belongs to. For each set of projected points, we learn a probability density function by a weighted sum of K Gaussian distributions:

$$p(x) = \sum_{k=1}^{K} \pi_k \; p(x|k) \tag{2}$$

with π_k being the weight of the k-th Gaussian and $p(x|k)$ being the conditional density function. The conditional density function is a d-dimensional Gaussian distribution:

$$p(x|k) = \frac{1}{\sqrt{2\pi}^d \; \sqrt{det(C_k)}} e^{-\frac{1}{2}\left((x-\mu_k)^T \; C_k^{-1}(x-\mu_k)\right)} \tag{3}$$

with mean μ_k and covariance matrix C_k. The above $p(x|k)$ can also be written as $\mathcal{N}(x|\mu_k, C_k)$. To estimate the parameters $\{\mu_k, C_k, \pi_k\}$ for each of the Gaussian kernels the Expectation-Maximization (EM) [4] algorithm is used. However, performing the EM algorithm in high dimensional spaces can be very time consuming. It is therefore convenient that our data is already projected to the low-dimensional PCA space as this ensures a fast convergence of EM.

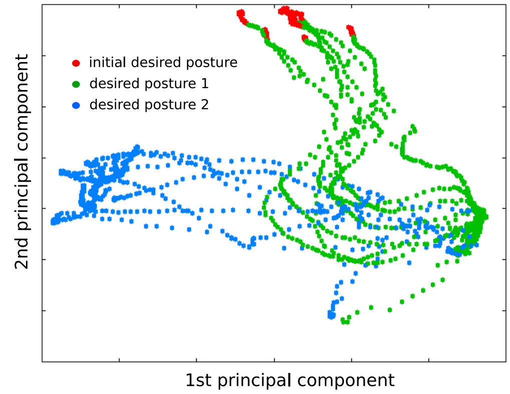

Fig. 3. Interaction data of several stand-up interactions projected into a low-dimensional posture space. Each point corresponds to one posture of the robot.

After learning, we end up with three GMMs coding three probability density functions $p_1(x), p_2(x), p_3(x)$. Each probability density function can be used to determine the probability of a point in low-dimensional posture space with respect to a particular desired posture. For example, computing $p_2(r)$ for a given projected robot posture r, returns the likelihood of the robot having desired posture $= 2$ when being in state r.

When the next interaction with the human is started, the robot can use the new learned model to decide in which state it is and which desired posture to take on. For this, the current joint values are projected onto the learned low-dimensional posture space. The result is a d-dimensional point. The optimal desired next posture can be computed in a maximum-likelihood fashion using:

$$s_{next} = \underset{s \in S}{argmax} \;\; p_s(x) \tag{4}$$

In each step of the control loop, the robot calculates s_{next} and sends the angular values of the desired posture to a low-level controller. The controller then computes the needed joint torques to take on this posture. After the interaction is finished, the human critique information is collected and used to update the memory. After that, the learning loop is repeated.

In Figure 3 we see an example of a set of interactions projected onto a low-dimensional space. Each point in the plot represents one posture of the CB^2 robot during an interaction. The points were colored according to the desired posture which was active in that particular time step.

IV. Experiment and Results

In order to investigate tightly coupled adaptation and the learning scheme proposed in this paper, we conducted an PHRI experiment using the rising-up interaction introduced earlier. In the experiment two subjects were asked to assist the robot in standing up. The first test-subject was part of

Fig. 4. Sequential snapshots of the first (top) and last (bottom) interaction of the test subjects with the robot. Left we see the expert user, right the beginner. The white curve depicts the change in the robot's hips position. The center figures of each sequential snapshot shows how the robot learns in both cases to have a strong contact between the feet and the ground.

the research team working on the CB^2 robot and, thus, often exposed to interactions with the robot. This test-subject will be referred to as *expert* subject in the following. The second user, has not been exposed to similar interactions in the past and will be referred to as *beginner*. The subjects had to repeatedly help the robot. After every 10 trials, the accumulated data in the memory was used for learning a new model, according to the learning scheme described in Section III-B. In total 30 interactions with 2 learning steps in between were conducted.

Figure 4 shows sequential snapshots of first and last interactions for each subject. The upper row of pictures shows the first interactions, while the lower row of pictures shows interactions after learning. The white dashed line indicates the height of the hips in each snapshot. In the figures we can observe a smoother transition of the hip height after the learning interaction, than before the learning interaction. Especially, in the center figures, we can see a strong contact between feet and the ground and an increased hip height after learning, in contrast to the poor contact with ungainly leg posture shown before learning. How much the human helped the robot in the task, and how the human evaluates the robot performance can be a subjective matter. Therefore, in our evaluation we focus only on whether the robot motion is refined such that inefficient and jerky motions are avoided.

Figure 5 shows three interactions for each user which were projected into the low-dimensional posture space. Each interaction is represented by a curve which reflects the robot's postures during the interaction. The black curve indicates the robot's postures during the initial phase, while the red and green curve indicate the robot's postures after the first and second learning step respectively. In the case of the expert user (left), we can clearly see that with each learning step, the interaction becomes smoother and shorter. In the initial phase the robot motion oscillates around the point $(0.5, -2.0)^T$. It might be caused by inefficient robot motion as discussed above. After each learning step, the robot motion becomes

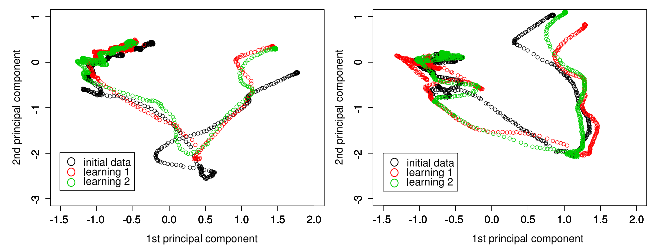

Fig. 5. Projected interactions in the low-dimensional posture space. Left: expert. Right: beginner.

smoother and more efficient, and the corresponding trajectory becomes smoother and shorter. In the diagram describing the beginner subject (right), this feature is not similarly obvious.

In order to confirm the above discussion we quantified the robot motion using the *posture change norm*. The posture change norm a of the robot was calculated using the Euclidean distance between the data of t and $t-1$ in the posture space $\boldsymbol{X}$ defined by using each joint angle as a base.

$$a_{(t)} = \| \boldsymbol{x}_{(t)} - \boldsymbol{x}_{(t-1)} \|_2, \ \boldsymbol{x} \in \boldsymbol{X}. \tag{5}$$

Computing the posture change norm at each time step of the interaction results in the time series depicted in Figure 6. The black time series' show the posture change norm during the initial interaction phase. We can see various sudden peaks indicating large changes in the robot posture and, thus, non-smooth motion. In particular in the case of the expert user, we can find high peaks (around 1000 msec). This is undesirable, as large changes in the robot posture result from strong forces acting on it. The green and red time series' show the evolution of the norm after each learning step. With each learning step, the amount and number of peaks in the time series is reduced. In other words, the fluctuations in the posture change norm decrease leading to a smoother and a more efficient motion. This supports the hypothesis, that the introduced learning method improves PHRI.

A statistical analysis of the data further underlines the above hypothesis. For this, we computed the sum and vari-

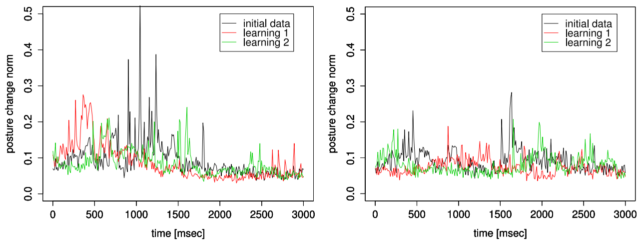

Fig. 6. Evolution of the posture change norm with increasing number of learning steps.

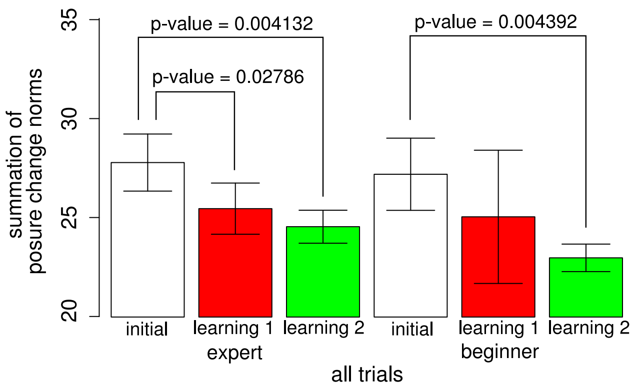

Fig. 7. The sum and stddev of the posture change norm of the expert and beginner subject. With each learning step, the amount of these two values decreases and the movement of the robot becomes smoother and more synchronized with that of the subject.

ance of the posture change norm during the interactions. Figure 7 shows the evolution of these values with each learning step. In both cases, we see that the sum and variance of the posture change norm decreased as the experiment progressed. These results are backed by a t-test confirming that the difference between the situation before and after learning is statistically significant.

In the light of the above results, an interesting question was raised: “Are the measured differences really due to a symmetric learning process, in which *both* human and robot adapt their behavior?”. In other words: does the robot learning system have any effect on the evolution of interaction? To investigate this question, we conducted a baseline experiment, in which we repeated the introduced experiment in slightly different setting. This time, learning in the robot was turned off, and fixed time steps were used for switching between the postures. Thus, the only kind of adaptation that is possible in this scenario, is the adaptation of the human towards the robot. Comparing these baseline results with the previously achieved results showed a significant difference, answering the above question, and affirming that the results achieved with the introduced probabilistic low-dimensional posture maps system are due to a bilateral learning process taking placing in *both* the human and the robot.

V. Conclusion

In this paper we presented a physical human-robot interaction scenario where successful task completion can only be achieved through coordinated actions involving physical contact. For this, we introduced a simple machine learning algorithm for adapting the behavior of the robot according to received critique from the human interaction partner. The method has a low computational load, can be run online while the interaction with the robot is going on and needs relatively few training data. In contrast to previous work in this field, the robot in this study is in close physical contact with the human partner and plays an active role during the execution of the cooperative task. The CB2 robot, through its flexible-joint design and soft silicone skin, is particularly suited for such tasks, as physical interactions become more ”natural” and lifelike. In an experiment inspired by the parenting behavior in humans, we were able to show that the proposed learning method results in measurable improvements of the interaction. Quantitative evaluations based on the posture change norm confirm the significance of these improvements.

In the future, we aim at investigating more complex interactive behaviors. Additionally, we hope to include an exploration phase into the learning algorithm, in which the robot can try out different variants of a behavior in order to find out, which one is best suited for the interaction partner. Further, we hope to include non-binary feedback from the human by using touch sensors or other input devices.

References

[1] E. Berger, H. Ben Amor, D. Vogt, and B. Jung. Towards a simulator for imitation learning with kinesthetic bootstrapping. In *in Proceedings of the SIMPAR 2008 International Conference on Simulation Modeling and Programming for Autonomous Robots, Workshop on The Universe of RoboCup Simulators, Proceedings CD*, 2008.

[2] S. Calinon and A. Billard. What is the teacher’s role in robot programming by demonstration? - Toward benchmarks for improved learning. *Interaction Studies. Special Issue on Psychological Benchmarks in Human-Robot Interaction*, 8(3):441–464, 2007.

[3] Agostino De Santis, Bruno Siciliano, Alessandro De Luca, and Antonio Bicchi. An atlas of physical human-robot interaction. *Mechanism and Machine Theory*, 43(3):253–270, March 2008.

[4] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum Likelihood from Incomplete Data via the EM Algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, 39(1):1–38, 1977.

[5] M. Hersch, F. Guenter, S. Calinon, and A. Billard. Dynamical system modulation for robot learning via kinesthetic demonstrations. *IEEE Trans. on Robotics*, 2008.

[6] Y. Hirata K. Kosuge, T. Hayashi and R. Tobiyama. Dance partner robot -ms dancer-. In *Proceedings of the 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2003.

[7] O. Khatib, K. Yokoi, O. Brock, K. Chang, and A. Casal. Robots in human environments. In *in Proceedings of the 1st Workshop on Robot Motion and Control*, pages 213–221, 1999.

[8] T. Minato, Y. Yoshikawa, T. Noda, S. Ikemoto, H. Ishiguro, and M. Asada. Cb2: A child robot with biomimetic body for cognitive developmental robotics. In *Proceedings of the 2003 IEEE-RAS/RSJ International Conference on Humanoid Robots. Humanoids ’07*, 2007.

[9] J. Tani, R. Nishimoto, J. Namikawa, and M. Ito. Codevelopmental learning between human and humanoid robot using dynamic neural network model. *IEEE Transactions on System, Man and Cybernetics*, 38(1):43–59, 2008.

[10] S. Yohanan and K. E. MacLean. The haptic creature project: Social human-robot interaction through affective touch. In *in Proceedings of the AISB 2008 Symposium on the Reign of Catz & Dogs: the Second AISB Symposium on the Role of Virtual Creatures in a Computerised Society*, volume 1, pages 7–11, 2008.