Physical Human–Robot Interaction

Mutual Learning and Adaptation

By Shuhei Ikemoto, Heni Ben Amor, Takashi Minato, Bernhard Jung, and Hiroshi Ishiguro

Close physical interaction between robots and humans is a particularly challenging aspect of robot development. For successful interaction and cooperation, the robot must have the ability to adapt its behavior to the human counterpart. Based on our earlier work, we present and evaluate a computationally efficient machine learning algorithm that is well suited for such close-contact interaction scenarios. We show that this algorithm helps to improve the quality of the interaction between a robot and a human caregiver. To this end, we present two human-in-the-loop learning scenarios that are inspired by human parenting behavior, namely, an assisted standing-up task and an assisted walking task.

Digital Object Identifier 10.1109/MRA.2011.2181676
Date of publication: 29 February 2012

Human–Robot Interaction and Cooperation

Until recently, robotic systems mostly remained in the realm of industrial applications and academic research. However, in recent years, robotics technology has significantly matured and produced highly realistic android robots. As a result of this ongoing process, the application domains of robots have slowly expanded into domestic environments, offices, and other human-inhabited locations. In turn, interaction and cooperation between humans and robots has become an increasingly important and, at the same time, challenging aspect of robot development. Particularly challenging is the physical interaction and cooperation between humans and robots. For such interaction to be successful and meaningful, the following technical difficulties need to be addressed:

- guaranteeing safety at all times
- ensuring that the robot reacts appropriately to the force applied by the human interaction partner
- improving the behavior of the robot using a machine-learning algorithm in a physical human–robot interaction (PHRI).

In our previous research [1], we presented a PHRI scenario in which we addressed the above topics. Inspired by the parenting behavior observed in humans, a test subject was asked to physically assist a state-of-the-art robot in a standing-up task. In such a situation, both the human and the robot are required to adapt their behaviors to cooperatively complete the task. However, most machine learning scenarios to date do not address the question of how learning can be achieved for tightly coupled, physical interactions between a learning agent and a human partner. Building on the results in [2], we present an extended evaluation and discussion of such human-in-the-loop learning scenarios.

To realize learning and adaptation on the robot's side, we employ a computationally efficient learning algorithm based on a dimensional reduction technique. In particular, after each trial, the human can judge whether the interaction was successful, then the judgment is used in a machine learning algorithm to apply a dimensional reduction technique and update the behavior of the robot. As learning progresses, the robot creates a behavioral model, which implicitly includes the actions of the human counterpart.

At the same time, refining the motions of the robot during a physical interaction requires the motions of the human to be improved, because the two motions influence each other. Hence, the human counterpart is part of the learning system and overall dynamics. To analyze the efficiency of the proposed learning algorithm and the effect of human habituation to the robot during such close-contact interactions, we perform a set of PHRI experiments. In addition to the assisted standing-up interaction scenario presented in [2], we also present and discuss the first results based on a novel interaction scenario. More specifically, we present an assisted walking task in which a human caregiver must assist a humanoid robot while walking.

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

Related Research

Important aspects of PHRIs have been investigated in a perspective research project conducted by the European Network of Excellence (EURON) [3]. The objective of the project was to present and discuss important requirements for safe and dependable robots involved in PHRIs. Initial approaches for achieving these requirements are currently being addressed in a follow-up research project called PHRIENDS (a PHRI that is dependable and safe). To reduce risks and fatalities in industrial manufacturing workplaces, the primary goal of the PHRIENDS project is to design robots that are intrinsically safe. This requires the development of new actuator concepts, safety measures, and control algorithms, 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 the focus of the PHRIENDS project.

Khatib et al. [4] discussed the basic capabilities needed to enable robots to operate in human-populated environments. In particular, they discussed how mobile robots can calculate collision-free paths and manipulate surrounding objects. In their approach, they characterized free space using an elastic strip approach. However, the described robots were not expected to come into direct (physical) contact with the surrounding human subjects. The importance of direct physical interaction was highlighted in the haptic creature project [5], which investigated the role of affective touch in fostering the companionship between humans and robots. In an attempt to improve human–robot interaction, Kosuge et al. presented a robot that can dance with a human by adaptively changing the dance steps according to the force/moment applied to the robot [6]. Amor et al. [7] used kinesthetic interactions to teach new behaviors to a small humanoid robot. Furthermore, the behavior of the robot may be optimized with respect to a given criterion in simulation. In this learning scheme, the robot is a purely passive interaction partner and acts only after the learning process is complete. Similar approaches to teaching new skills have also been reported in [8] and [9] using different learning methods, i.e., continuous time-recurrent neural networks and Gaussian mixture models (GMMs), respectively. Odashima et al. [10] developed a robot that can come into direct physical contact with humans. This robot is intended for caregiving tasks such as carrying injured persons to a nearby physician. The robot can also learn new behaviors and assistive tasks by observing human experts as they perform these tasks. However, this learning does not take place during interactions but rather in offline sessions using immersive virtual environments. In [11], Evrard et al. present a humanoid robot with the ability to perform a collaborative manipulation task together with a human operator. In a teaching phase, the robot is first teleoperated using a force-feedback device. The recorded forces and positions are then used to learn a controller for the collaborative task. The main hypothesis underlying this approach is that the intentions of the human interaction partner can be guessed from haptic cues. In [12], physical interactions between a robot's hand and the hand of a human are modeled by recording their distances. The distances are then encoded in a hidden Markov model (HMM), which in turn is used to synthesize similar hand contacts. A recent survey on modern approaches to physical and tactile human–robot interaction can be found in [13].

In this article, we present experiments with a flexible-joint 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 to learn and adapt their behaviors to their partner so as to achieve a common goal. This tight coupling of robot and human learning and coadaptation is a unique feature and is the primary contribution of the present study. We assume that it is important to focus on the active role in the interaction because the forces generated during the active behavior of the robot influence the behavior of the human, which in turn influences the passive behavior of the robot. In addition, these active and passive roles cannot be clearly separated because the robot and the human influence each other when they are in physical contact.

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. Figure 1 shows an overview of the learning scheme used in this article. 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 a success or failure (binary evaluation). The feedback can be provided in various ways, such as through touch or through a simple graphical user interface. Once the evaluation information is collected by the robot system, it is stored in a database in the memory. The memory collects information about recent successful interactions and manages the data for the subsequent learning step. This allows us to optimize the set of training examples used for learning to improve learning quality. Figure 1 shows the human-in-the-loop learning system considered in this article, where the behavior of the human influences the behavior of the robot and, simultaneously, the behavior of the robot influences the behavior of the human. Furthermore, the behavior of the robot changes as learning progresses, which in turn influences the behavior of the human and its physical support. This system demonstrates one of the applications of a tightly coupled physical interaction.

After a number of interactions, the learning system queries the memory for a new set of training data. The data are then projected onto a low-dimensional manifold using dimensional reduction techniques. There are three justifications for this step. First, dimensional reduction allows a reduction of the space in which learning takes place. Thus, the learning can be much faster and more efficient. In addition, 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, dimensional reduction reduces the negative influence of outliers on learning. The inputs to the dimensional reduction step are high-dimensional state vectors describing the postures of the robot during the interaction. The output is a low-dimensional posture space.

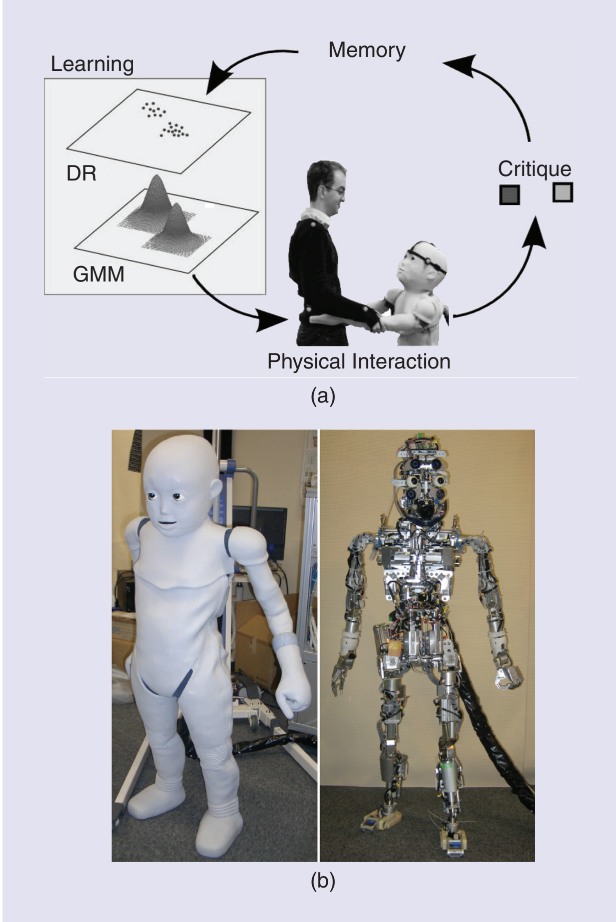

Figure 1. (a) Overview of the physical interaction learning approach. After physical interaction, the human judges whether the interaction was successful. This information is stored in the robot's memory and used for later learning. (b) Flexible-joint humanoid robot used in the experiments in this study. (Photos courtesy of ERATO Asada Project.)

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 obtain for each possible action a set of states in which the corresponding action should be triggered. For each action, a GMM is learned. The model encodes a probability density function of the learned state vectors. The ideal number of Gaussian mixtures is estimated using the Bayesian information criterion (BIC) [14].

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 actions of the robot. Here, 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 executed by the robot.

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

The CB^2 Robot

The robot used in this study is called the child–robot with biomimetic body, or CB^2 [15]. The robot has the following features.

- Its height is 130 cm, and its mass is approximately 33 kg.
- The degree of freedom (DOF) is 56.
- The supplied air pressure is 0.6 MPa.
- The efficient torque of the knee is theoretically 28.6 N·m.
- 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, are equipped with potentiometers.

The joints have low mechanical impedance because of the compressibility of air. The joints can also be made to be completely passive if the system discontinues air compression during robot motion. This helps the robot to perform passive motions during physical interaction and helps to ensure the safety of the human partner. This is in contrast to most other robots, in which the joints are driven by electric motors with decelerators. The flexible actuators enable the joints to produce seemingly smooth motions, even when the input signal changes drastically. This feature of the CB^2 robot is used to realize complex motions using the simple control architecture [1] depicted in Figure 2. More specifically, full body motions of the robot are realized by switching between a set of successive desired postures. Furthermore, the flexible actuators enable motions generated by this simple control architecture to be adaptively changed in response to an applied force from the human partner. Each posture is described by a posture vector x, with each entry of the vector denoting the angular value of a particular joint. A low-level controller is implemented by the proportional-integral-differential (PID) control of angular values. Each time the desired posture is switched drastically, large drive torques are generated, resulting in an active force being applied to the human caregiver. As the posture of the robot approaches the desired posture, the passive motion gradually becomes the dominant motion of the robot because the amount of error in the angular control gradually becomes smaller.

Figure 3 shows how the examined standing-up task is realized using the proposed control architecture. The behavior is realized by switching between three desired postures. At first glance, the specifications of the robot motion appear to be extremely simple. However, the switching times are highly dependent on the human interaction. 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 the characteristics of its partner during the period of interaction. In addition, it must be noted that this motion cannot be performed by the robot if a human does not assist in its execution.

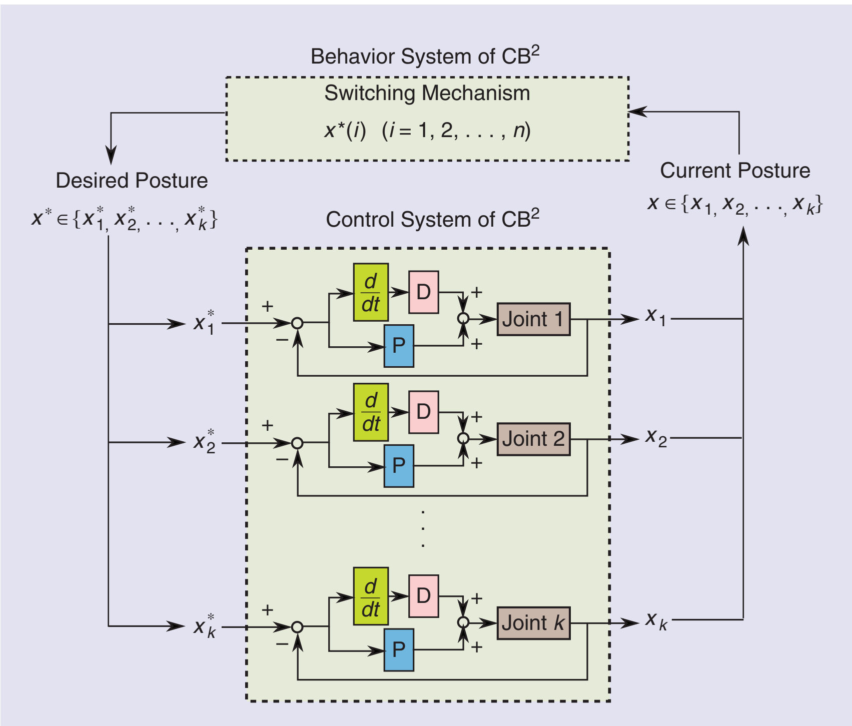

Figure 2. Control architecture of the CB^2 robot. The desired posture is encoded as a vector $\mathbf{x}^*$ of angular values. Using a PID controller, drive torques are generated to attain the desired posture. The switching mechanism changes between a set of different desired postures to achieve complex robot motions.

Learning Method

In the standing-up task, the goal of learning is to determine the ideal timing for switching actions $x^* \in X^* \subseteq X$ between different desired postures. Here, x^* is a desired posture, X^* is a set of desired postures prepared for control, and X is a posture space that is constructed from all joint angles. This is achieved by learning three different probabilistic low-dimensional posture models: 1) for

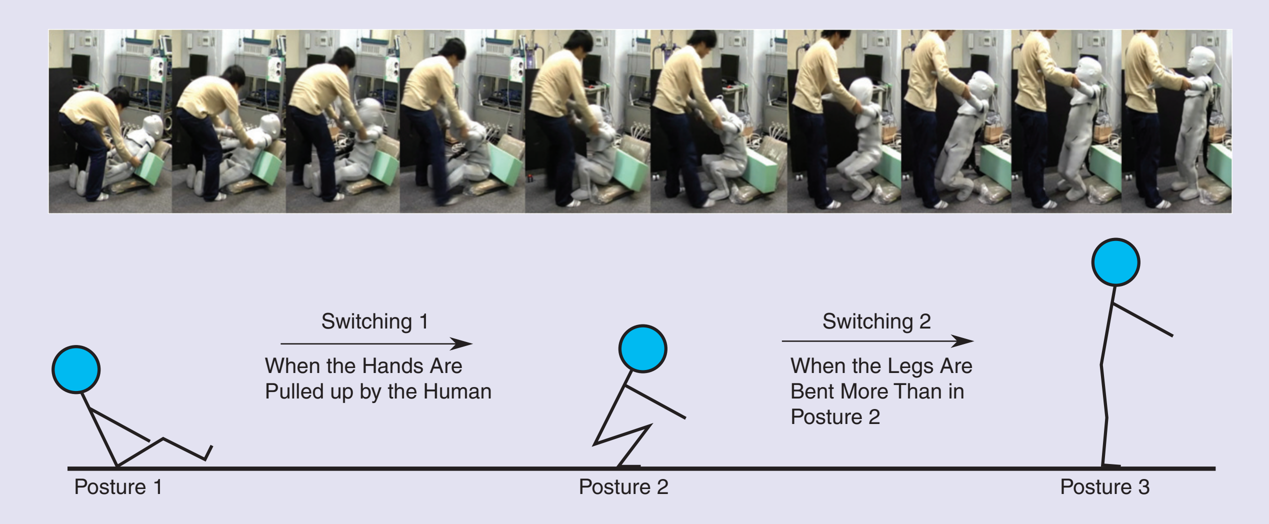

Figure 3. The three desired postures used in the standing-up task of the experiment. The learning task is to determine the ideal switching conditions between the desired postures. (Photo courtesy of ERATO Asada Project.)

the case in which no switching occurs, 2) for the first switching action, and 3) for the second switching action.

At each time step of an interaction between the human and the robot, the realized posture and the current desired posture of the robot are recorded. The robot posture r is a 52-dimensional vector that codes the current angular value of each joint. After the interaction is complete, the postures are stored in a database in the memory. The database holds the information for the last ten interactions. Although there are several possible ways to integrate this new data into the database, the general policy used here is this new data overwrite old data, and successful interactions overwrite failed interactions.

After ten interactions, the training data from the memory are used for learning. The goal of the learning is to construct a model that indicates when the robot should switch actions by changing the current desired posture. This rule is described by a mapping from the current posture of the robot to the desired posture that the robot should use. To realize this map, we use a GMM that can construct a probabilistic model. Therefore, the objective model of the learning is a probabilistic model that indicates the likelihood of desired postures in the current state.

First, dimensional reduction is applied to the data because a 52-dimensional vector has too many dimensions to learn the model. Although a number of methods can be applied for this task, in this article, we used a principal component analysis (PCA). To perform the 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, also known as the eigenvectors or principal components (PCs), of matrix M. The matrix W is a diagonal matrix containing singular values. Each PC has a corresponding singular value that indicates how much information of the data set is covered by a specific PC. The first few PCs are then used as the axes of the 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 PCs.

Next, we compute a GMM for each of the three switching classes. Here, we divide the projected data points into distinct sets. If no switching occurred, then the corresponding point is assigned to the first data set. Otherwise, the corresponding point is assigned to one of the other two sets. 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 kth 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 $\mathcal{N}(x|\mu_k, C_k)$. The expectation-maximization (EM) [16] algorithm is used to estimate the parameters $\{\mu_k, C_k, \pi_k\}$ for each of the Gaussian kernels. Fortunately, performing the EM algorithm in low-dimensional spaces improves the convergence of the algorithm.

After the learning process, we end up with three GMMs coding three probability density functions, namely, $p_1(x)$, $p_2(x)$, and $p_3(x)$. In our experiments, each GMM had between five and ten Gaussians. Each probability density function can be used to determine the probability of a point in a low-dimensional posture space with respect to a particular switching action. For example, computing $p_2(r)$ for a given projected robot posture, r, returns the likelihood of the robot having to switch from the second to the third desired posture when the robot is in state r.

When the next interaction with the human starts, the robot can use the newly learned model to decide its current state and the desired posture. Here, the current joint values are projected onto the learned low-dimensional posture space. The result is a d-dimensional point. The optimal desired subsequent switching action can be computed in a maximum-likelihood fashion as follows:

$$x^*_{\text{next}} = \underset{x^* \in X^*}{\operatorname{argmax}}\, p_s(x). \tag{4}$$

In each step of the control loop, the robot calculates s_{next} and sends the angular values of the corresponding desired posture to a low-level controller. The controller then computes the needed joint torques to take on this posture. After the interaction is complete, the human evaluation information is collected and used to update the memory. The learning loop is then repeated. The above algorithm is closely related to HMMs [17]. At the same time, however, our algorithm deviates in various ways from HMM. More specifically, we do not learn the sequencing of states in our system. As a result, no explicit transition probabilities between the states are modeled.

Figure 4 shows 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 that was active during that particular time step.

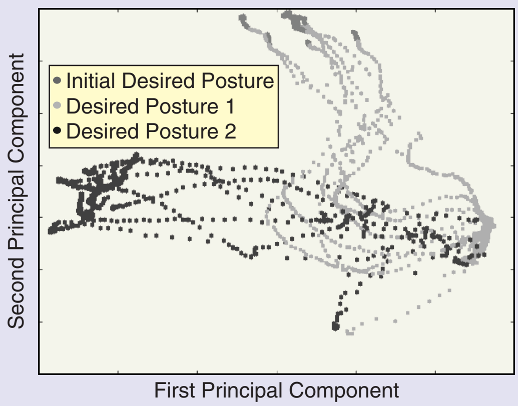

Figure 4. Interaction data for the standing-up task projected into a low-dimensional posture space. Each point corresponds to one posture of the robot.

Experiment and Results

To investigate tightly coupled adaptation and the learning scheme proposed in this article, we conducted a PHRI experiment using the interaction for the standing-up task introduced earlier. In particular, we considered the following question: "Does the learning algorithm lead to a symmetric learning process, in which both human and robot adapt their behaviors?" Furthermore, we wanted to measure the contribution of the learning algorithm to any improvement in the interaction. This required a careful experiment design that would allow us to distinguish between learning-based adaptation and adaptation due to human habituation to the robot.

The experiment was split into three independent parts. Throughout the experiment, five subjects were asked to repeatedly assist the robot in standing up. In the first part, after every ten trials, the accumulated data in the memory were used for learning a new model, according to the learning scheme described in the "Physical Interaction Learning Approach: Learning Method" section. In total, 30 interactions with two intermediate learning steps were performed. In the second part of the experiment, learning by the robot was disabled and fixed time steps were used for switching between the postures. In this baseline scenario, the only type of adaptation that was possible was the adaptation of the human to the robot. In the third and final part, learning was once again enabled (the results of the first part were not included; hence, learning started from the beginning again). The experimental design ensures that we have baseline data, allowing us to compare the results of the interactions with and without learning. In addition, by performing the baseline experiment between the learning experiments, we ensure that the user is already familiar with the robot. Thus, we rule out any distortion of the baseline result because of unfamiliarity.

To determine the ideal number of PCs on which to project the 52-dimensional posture vector of the robot, intrinsic dimensionality estimation techniques can be used [18] as a criterion. A simple estimation technique is based on the analysis of eigenvalues, which store the amount of information that is captured by each of the PCs. Hence, the eigenvalues determine how many PCs are needed to retain a specific percentage of information found in the data set. In our implementation, we automatically determine the number of PCs that capture more than 85% of the information in the data set. For our standing-up data set, this resulted in a projection onto two PCs.

Figure 5 shows sequential photographs of the interactions of two test subjects. Figure 5(a) shows the initial interaction, whereas Figure 5(b) shows the interaction 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, when compared with that before the learning interaction. In particular, the center photographs reveal strong contact between the feet and the ground and an increased hip height after learning, in contrast to the poor contact with the ungainly leg posture beforehand. Since the degree to which the human helped the robot in the task and the evaluation of the robot performance are somewhat subjective, in our evaluation, we focus only on whether the robot motion is refined to the degree that inefficient and jerky motions are avoided.

Figure 6 shows the interaction trajectories for two users before and after learning. Each trajectory was computed by projecting the robot postures into the low-dimensional posture space. Before learning, the trajectories contain loops and are partially linear. These linear pieces of the trajectories are due to jerky movements and large changes in the robot postures. In particular, for the first user, the variance in the trajectory decreases after learning. The trajectories become more similar and take on a V-shaped form. This can be explained by the fact that the interaction consists of three desired postures. Therefore, in successful trials, the interaction leads the robot from a starting posture to an intermediate posture and then to a final posture, as shown in Figure 3. In a low-dimensional space, the result is a V-shaped or triangular-shaped trajectory. This allows us to qualitatively evaluate the efficiency and naturalness of the interaction by analyzing the smoothness and shape of the low-dimensional trajectories. For example, in the case of the second subject, the trajectories before learning contain large loops at the point $(1.7, -1.5)^T$, which is the low-dimensional coordinate of the second desired posture. This phenomenon can easily be explained if we take into account our previous analysis. In the initial trials, the robot has poor contact with the floor and the legs are often not symmetrically arranged when reaching the second desired posture. As a result, lifting the robot becomes more difficult for the human and involves slight modifications of the robot posture to make the feet more stable. This interrupts the flow of the standing-up task and increases the interaction burden for the human caregiver.

To confirm the above discussion, we quantified the robot motion using the posture change norm. The posture change norm a of the robot motion was calculated using the Euclidean distance between the data of t and $t-1$ in the posture space X defined using each joint angle as a base:

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

Computing the posture change norm at each time step of the interaction results in the time series depicted in Figure 7. The solid line shows the posture change norm during the initial interaction phase. We can see a sudden peak indicating a large change in the robot posture and, consequently, a nonsmooth motion. This is undesirable because large changes in the robot posture result from strong forces acting on the robot. The other lines show the evolution of the norm after each learning step. With each learning step, the 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 more efficient motion.

A statistical analysis of the data further underlines the above hypothesis. Here, we computed the mean and standard deviation of the summation of the posture change norm during the interactions. Figure 8 shows the evolution of these values with each learning step. For all subjects, we see that the mean and standard deviation of the posture change norm decreased as the experiment progressed. In the baseline experiment, only one subject was able to significantly improve the interactions, where statistical significance is computed using a t test. None of the other subjects

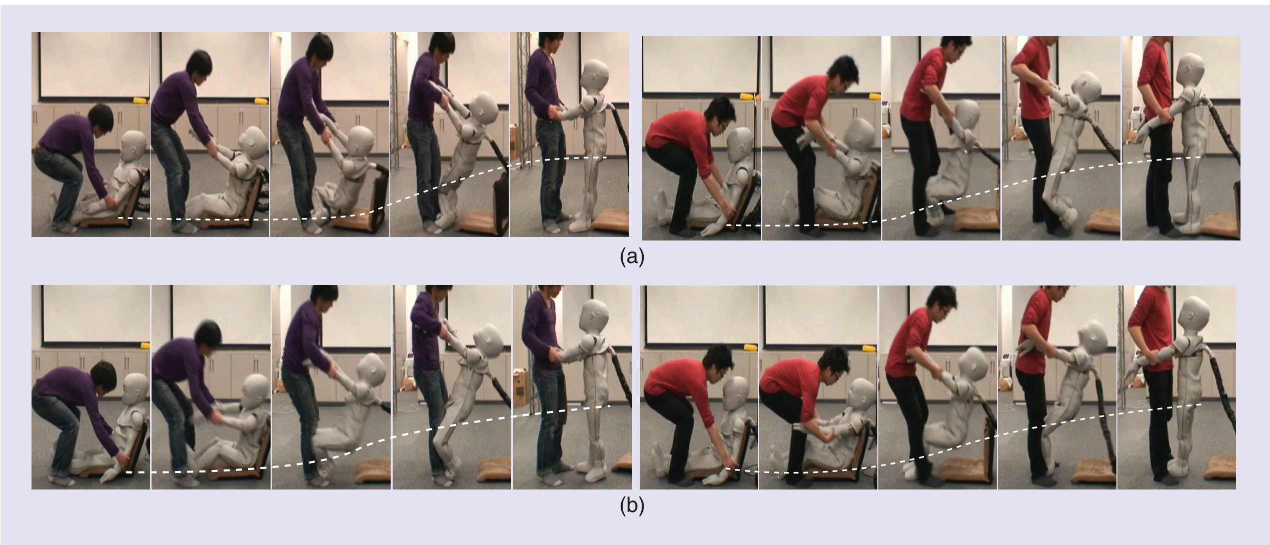

Figure 5. Sequential photographs of the (a) first and (b) last interactions of the test subjects with the robot. The white curve depicts the change in position of the robot's hips. The center photograph of each sequence shows how the robot learns to maintain firm contact between its feet and the ground for both subjects. (Photos courtesy of ERATO Asada Project.)

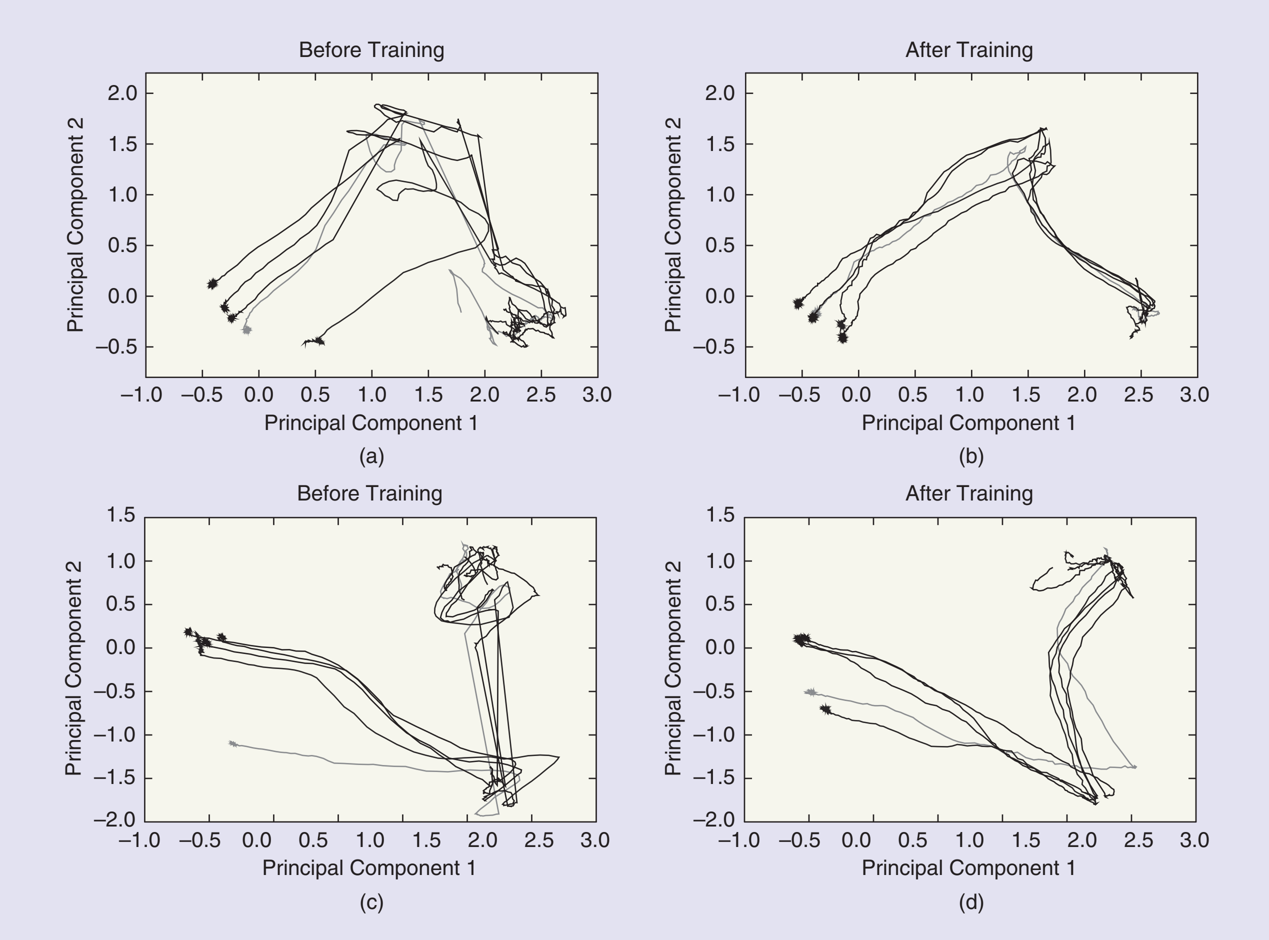

Figure 6. Projected interactions in the low-dimensional posture space: (a) and (b) the interaction trajectories for the first subject before and after learning and (c) and (d) the interaction trajectories for the second subject. In both the cases, the trajectories become smoother after learning and sudden jumps and knots are reduced. Furthermore, the trajectories become V-shaped, clearly indicating a smooth transition between the three desired postures.

were able to improve their interactions. In the first experiment, in which the proposed learning system is used, three subjects show significant improvement. Finally, in the second learning experiment, all of the subjects showed significant improvement in their interactions. This indicates that while a human can adapt to a robot and thus improve their interactions (as in the baseline experiment), this adaptation can be significantly improved by empowering the robot with learning capabilities (first and second learning experiments). We also analyzed the maximum values of the posture change norm during the interaction. Figure 9 shows the change in the maximum posture change norm during each learning phase of the baseline experiment and the first learning experiment. No significant difference in the maximum posture change norm is observed in the baseline experiment. On the other hand, in the learning experiment, there are large changes in the maximum posture change norm. For all subjects, the values drastically decrease after learning.

Still, one possible implication from above results cannot be ruled out by the experiments performed so far. Specifically, it remains unclear how much the learning system contributes to the improvement of interaction. A possible argument would be that the observed improvements are due to the long-term habituation and experience with the robot. If this argument is true, then we should see a similar improvement of interactions as above, even if we simply repeat the baseline experiment (where learning is disabled) three times in a row. To investigate this question, we performed the aforementioned experiment (three times baseline) with all subjects. For the subjects, the experiment looked exactly the same as the other experiments: the difference was not transparent. Figure 10 compares the summation of posture change norm between the first and the third baseline experiment. In each of the experiments, only one subject made significant improvement during the intermediate learning steps. On the whole, although for some subjects slight improvement was visible (notably Subject 5), the results are not as comprehensive as when learning is enabled. This means that, while long-term habituation and experience aids the learning process, it is not sufficient for a general improvement in PHRI.

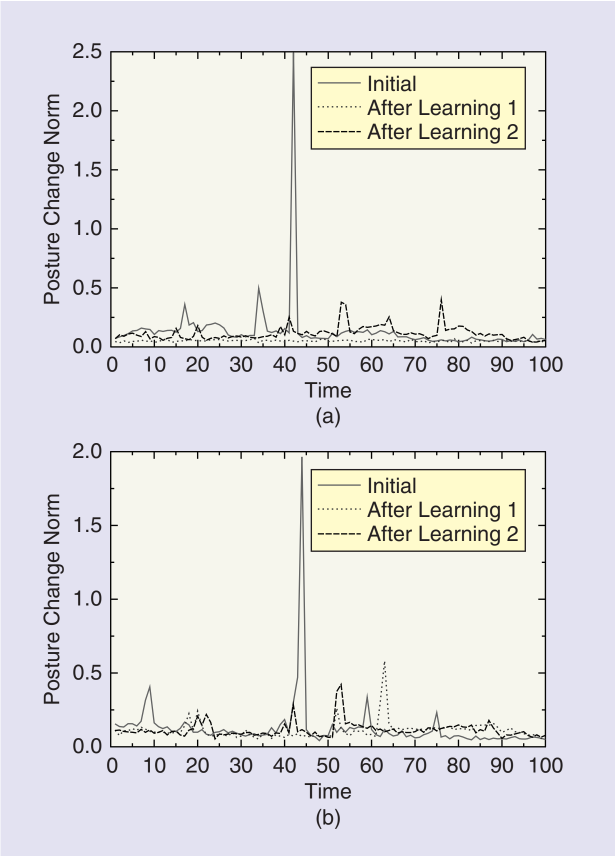

Figure 7. Evolution of the posture change norm during one learning experiment. The solid, dotted, and dashed lines show the evolution of the value when the robot has not yet learned, after the first intermediate learning step, and after the second intermediate learning step, respectively. (a) Subject 1 and (b) Subject 2.

Discussion

The following observations are based on the results of the above experiments. First, when learning and adaptation were only possible on the side of the human caregiver, generally, little or no improvement could be measured. However, even in this asymmetric learning situation, at least one subject was able to adapt to the robot so as to significantly improve the interaction quality. This shows the human ability to quickly adapt to new situations and motor tasks. The second observation is that the interaction quality significantly improved in the first learning experiment, and the improvement was even more remarkable during the second learning experiment. These results support our working hypothesis that the proposed learning system facilitates PHRI. Another interesting observation is that the human adaptation to the robot occurred in stages throughout the experiment. At the beginning of the experiment, the users were intimidated by the robot and the experimental setup. However, during the course of the experiment, the test subjects became more and more comfortable with the situation and the robot dynamics. As a result, the test subjects found it easier to interact with the robot. This suggests that algorithms for improving PHRI

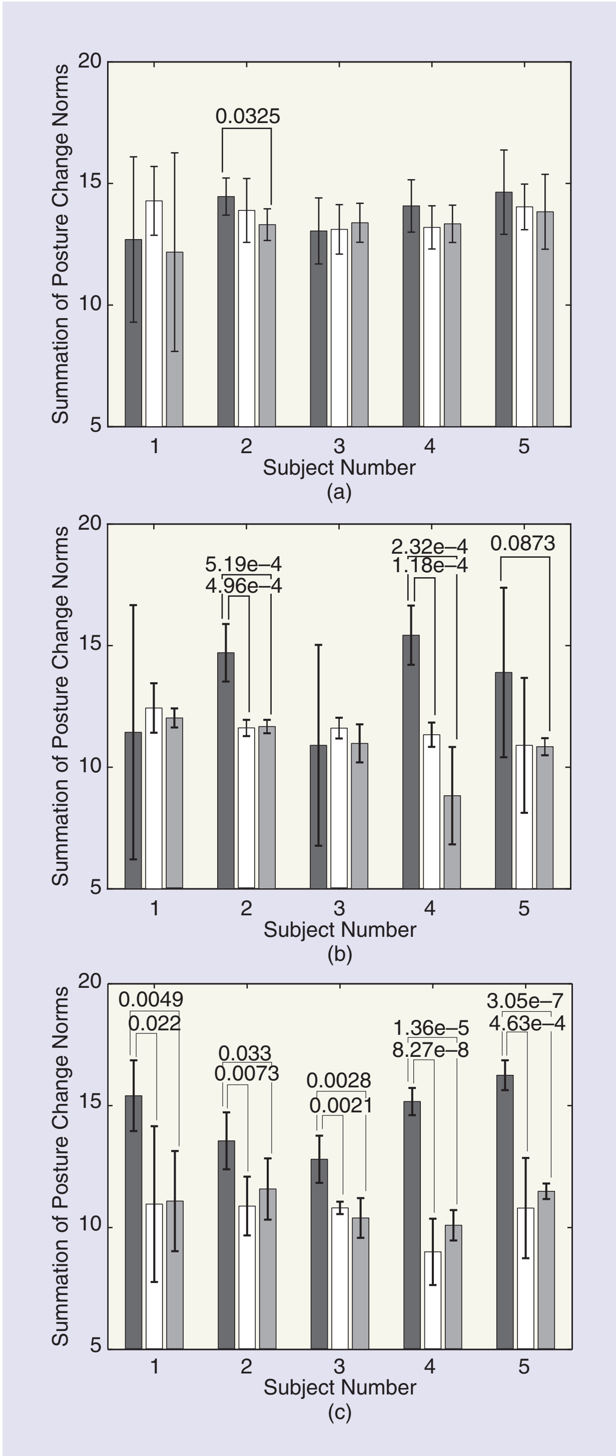

Figure 8. Mean and standard deviation of the summation of the posture change norm of test subjects in the (a) baseline, (b) first learning, and (c) final training experiments. The dark gray, white, and light gray bars indicate the mean and standard deviation values during each of the intermediate learning steps (after every ten trials). In (a), the baseline experiment, only Subject 2 shows a significant improvement after all trials. In (b) the first learning experiment, Subjects 2, 4, and 5 show significant improvements. In (c) the final experiment, the interaction with the robot improved for all subjects. With each learning trial, the indicated values decrease, and the movement of the robot becomes smoother and more synchronized with that of the subject.

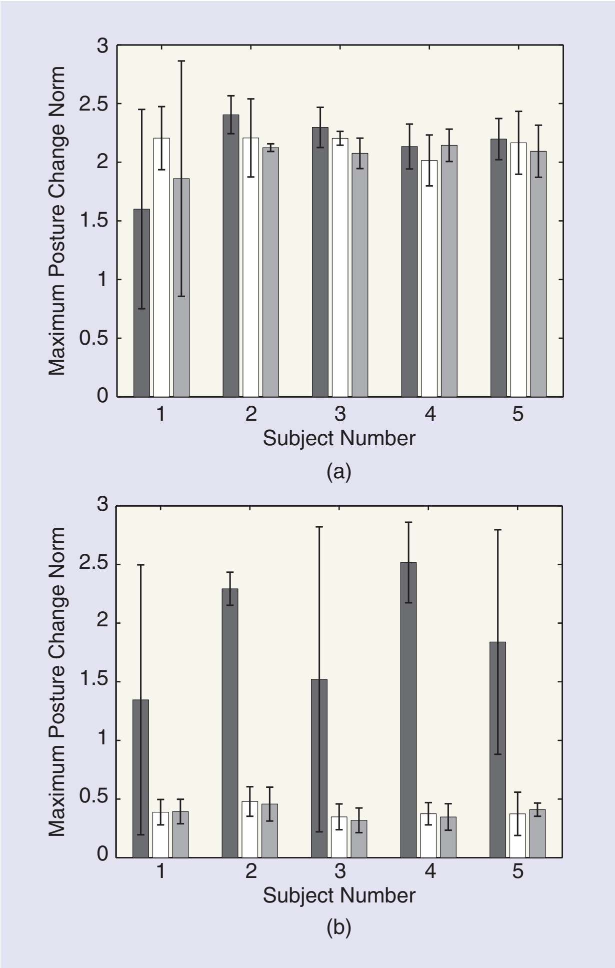

Figure 9. Change in the maximum posture change norm in each phase during (a) the baseline and (b) the first learning experiments. No significant difference in the maximum posture change norm is observed in the baseline experiment. On the other hand, there are large changes in the maximum posture change norm in the learning experiment. For all subjects, the values decrease drastically after learning.

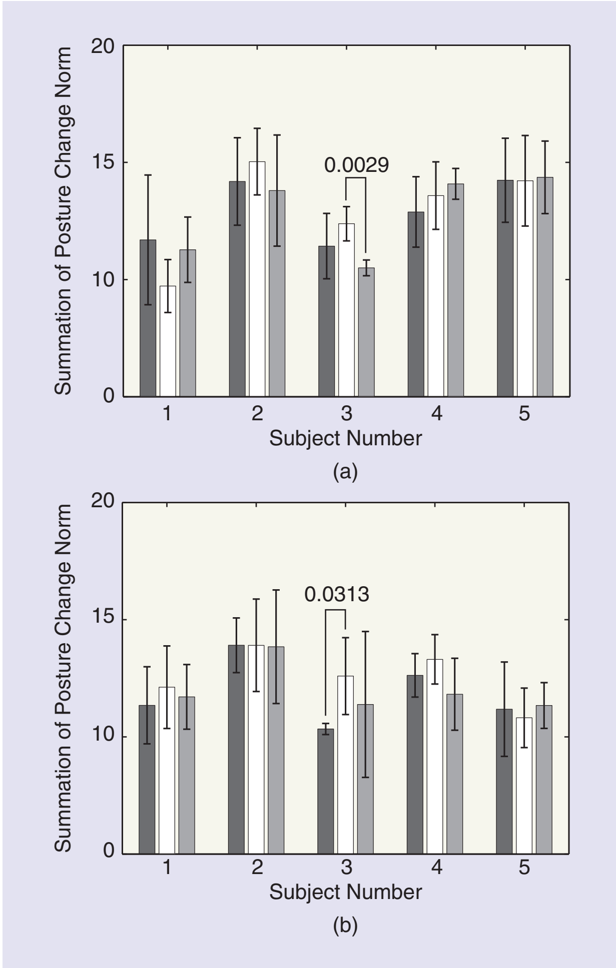

Figure 10. Baseline experiment repeated three times in a row to investigate whether improvement can be made without the robot's learning system enabled. In each of the experiments only one subject made a significant improvement. On the whole, although for some subjects slight improvement is visible (notably Subject 5), the results are not as comprehensive as when learning is enabled. (a) First baseline and (b) third baseline experiments.

can be made more efficient if the familiarization of the human with the robot is taken into account. A special familiarization phase, in which the human caregiver becomes accustomed to the robot before any cooperative tasks, might be one approach. Another method by which to familiarize the human with the robot might be a well-designed interaction protocol that involves tasks that are intended only to familiarize the human with the robot. An interesting feature of the proposed algorithm is the ability to monitor the progress of learning as trajectories in a low-dimensional space. The results of this study indicate that the trajectories converge toward a V-shaped pattern for the standing-up task. Furthermore, the trajectories, after learning, appear to have particular points or bottlenecks through which they pass. This is reminiscent of the study by Kuniyoshi et al. [19] in which it was shown that the dynamic motions for a particular task often have a bottleneck in the state space. This bottleneck is the result of the interaction of the human body and the environment. Kuniyoshi et al. referred to this property as knack and showed that the knack can be exploited to efficiently control a humanoid robot. In the proposed PHRI scenario, the dynamics of the robot strongly depends on the dynamics of the human caregiver. A knack may be said to appear in PHRI because of the strong coupling between the human and the robot and the resulting joint dynamics. In other words, the human can be regarded as a changing environment that constraints the robot dynamics. Note that, although only the posture of the robot was used to create the trajectories, we can still discern a knack that is based on joint dynamics. However, it can be argued that posture information is not sufficient enough to draw final conclusions about the joint dynamics. To address this question, we are currently investigating a different cooperative

PHRI task, namely that of assisted walking as can be seen in Figure 11.

In this scenario, the human caregiver must assist the robot while the latter is trying to walk. Similar to the standing-up task, the assisted walking is realized using three desired postures: left leg up, standing, and right leg up. These postures are repeated in a predetermined order (standing → left leg up → standing → right leg up) to create a cyclic walking motion. During an interaction session, the human assists the robot in performing four cycles of the latter sequence. For a fast assessment of the applicability of our approach to different scenarios, we performed an experiment using the same setup and parameters as for the standing-up task. However, in this case, we had only one test subject performing 30 interactions with learning enabled and 30 interactions as baseline. Figure 12 shows the comparison of posture change norms between each phase (a phase consists of ten trials) in each experiment (one baseline and one learning experiment). As opposed to the baseline experiment, we can see that the posture change norms decrease when learning is enabled. Note that the baseline experiment was performed after the learning experiment to account for the human's habituation.

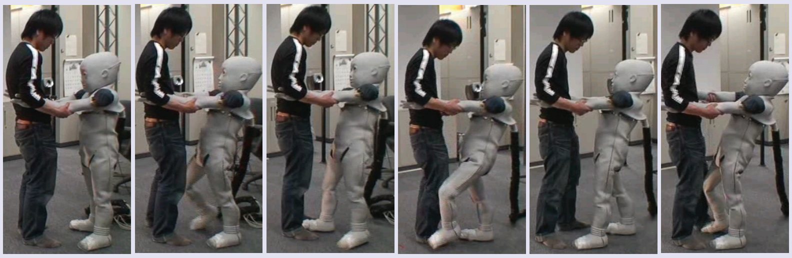

Figure 11. The assisted walking task where a human caregiver assists the robot in his or her attempt to perform several walking steps. (Photo courtesy of ERATO Asada Project.)

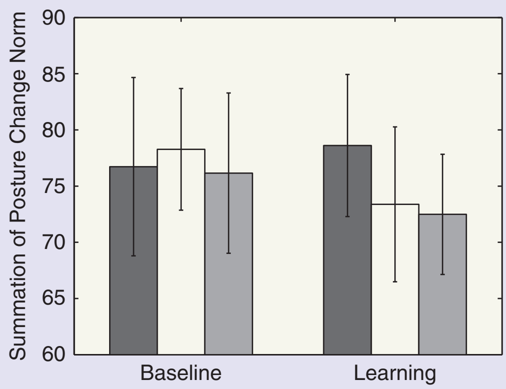

Figure 12. Summation of posture change norm for baseline and learning experiments in the assisted walking task. Each bar corresponds to a phase of ten interactions with the robot.

These early results show that the proposed human-in-the-loop learning system is not limited to the uprising interaction and that other types of interactions can be realized. At the same time, in our experiments, we found that the robot often failed to keep up when the human demonstrator drastically increased or reduced the speed of his or her walking gaits. This is due to the reactive nature of estimating the joint dynamics from the postures only. To keep up with a human interaction partner in this scenario, the robot must be more predictive in its estimation of the joint dynamics. One possible approach to overcome this problem is to include sensor information into the probabilistic low-dimensional posture models. That is, the state of the robot would be based on the current joint angles as well as the information gathered from the sensors under the skin. In this case, switching between one posture and another would also be influenced by the amount of pressure exerted by the human caregiver on the robot's body, e.g., the arms during assisted walking. Further studies are underway to obtain a conclusive answer to these questions.

Conclusions

In this article, we presented a PHRI scenario in which successful task completion can only be achieved through coordinated actions involving physical contact. We introduced a simple machine learning algorithm for adapting the behavior of the robot according to an evaluation by a human interaction partner. This method has a low computational load and can be run online during the interaction with the robot and requires relatively few training data. In contrast to previous research in this field, the robot considered in this study is in close physical contact with the human partner and plays an active role during the performance of the cooperative task. The CB^2 robot, through its flexible-joint design and soft silicone skin, is particularly well suited to such tasks because physical interactions become more natural and lifelike. In an experiment inspired by parenting behavior in humans, we were able to show that the proposed learning method results in measurable improvements of interaction. Quantitative evaluations based on the posture change norm confirm the significance of these improvements.

Thus far, the control system used herein has three parameters: the set of desired postures, the feedback gains, and the switching rule. In this article, we focused on learning the switching rule only. However, for more complex interaction scenarios it might be important to adapt all of these parameters. Another limitation of the proposed learning algorithm is the use of binary evaluation information. As a result, optimization of the parameters in a gradient descent manner is not possible. Another drawback of binary evaluation information is that only positive feedback examples are retained for use in the learning set while negative feedback examples are removed from the learning set. With respect to the first limitation, the desired postures and feedback gains can be regarded as attractors and velocities in a low-dimensional space. Amor et al. [7] have shown that such attractors can be efficiently learned in a low-dimensional space while incorporating kinesthetic assistance provided by the user. In the future, we therefore hope to integrate such a method into the proposed PHRI algorithm. As for the second limitation, we are considering the use of pressure sensors on the body of the robot. The amount of pressure issued by the caregiver can then be used as an approximate evaluation information. This allows for a finer grained reward value and, consequently, the use of modern optimization algorithms. Pressure sensors are also helpful to distinguish whether the human is currently in contact with the robot.

In summary, this study provided interesting insights into the dynamics of PHRIs. The combination of a soft-body robot and an efficient learning scheme is an important step toward responsive robots that share a common living space with humans.

References

[1] S. Ikemoto, T. Minato, and H. Ishiguro, "Analysis of physical human-robot interaction for motor learning with physical help," *Appl. Bionics Biomech. (Special Issue on Humanoid Robots)*, vol. 5, no. 4, pp. 213–223, 2008.

[2] S. Ikemoto, H. B. Amor, T. Minato, H. Ishiguro, and B. Jung, "Physical interaction learning: Behavior adaptation in cooperative human-robot tasks involving physical contact," in *Proc. IEEE Int. Symp. Robot and Human Interactive Communication (Ro-Man)*, Sept. 2009, pp. 504–509.

[3] A. De Santis, B. Siciliano, A. De Luca, and A. Bicchi, "An atlas of physical human-robot interaction," *Mechanism Mach. Theory*, vol. 43, no. 3, pp. 253–270, Mar. 2008.

[4] O. Khatib, K. Yokoi, O. Brock, K. Chang, and A. Casal, "Robots in human environments," in *Proc. 1st Workshop Robot Motion and Control*, 1999, pp. 213–221.

[5] S. Yohanan and K. E. MacLean, "The haptic creature project: Social human-robot interaction through affective touch," in *Proc. AISB Symp. Reign of Catz & Dogs: The 2nd AISB Symp. Role of Virtual Creatures in a Computerised Society*, 2008, vol. 1, pp. 7–11.

[6] K. Kosuge, T. Hayashi, Y. Hirata, and R. Tobiyama, "Dance partner robot—MS-danSer," in *Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems*, 2003, vol. 3, pp. 3459–3464.

[7] H. B. Amor, E. Berger, D. Vogt, and B. Jung, "Kinesthetic bootstrapping: Teaching motor skills to humanoid robots through physical interaction," *KI 2009: Advances in Artificial Intelligence* (ser. Lecture Notes in Artificial Intelligence), B. Mertsching, Ed. Berlin, Germany: Springer-Verlag, 2009, pp. 492–499.

[8] J. Tani, R. Nishimoto, J. Namikawa, and M. Ito, "Codevelopmental learning between human and humanoid robot using dynamic neural network model," *IEEE Trans. Syst., Man, Cybern.*, vol. 38, no. 1, pp. 43–59, 2008.

[9] 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)*, vol. 8, no. 3, pp. 441–464, 2007.

[10] T. Odashima, M. Onishi, K. Tahara, K. Takagi, F. Asano, Y. Kato, H. Nakashima, Y. Kobayashi, T. Mukai, Z. W. Luo, and S. Hosoe, "A soft human-interactive robot ri-man," in *Proc. IEEE/RSJ Int. Conf. Intelligent Robotics and Systems (IROS, v018)*, 2006, p. 1–1 (video session).

[11] P. Evrard, E. Gribovskaya, S. Calinon, A. Billard, and A. Kheddar, "Teaching physical collaborative tasks: Object-lifting case study with a humanoid," in *Proc. IEEE-RAS Int. Conf. Humanoid Robots (Humanoids)*, Dec. 2009, pp. 399–404.

[12] D. Lee, C. Ott, and Y. Nakamura, "Mimetic communication with impedance control for physical human-robot interaction," in *Proc. IEEE Int. Conf. Robotics and Automation (ICRA)*, 2009, pp. 1535–1542.

[13] B. D. Argall and A. G. Billard, "A survey of tactile human-robot interactions," *Robot. Autonom. Syst.*, vol. 58, no. 10, pp. 1159–1176, 2010.

[14] G. Schwarz, "Estimating the dimension of a model," *Ann. Statist*, vol. 6, no. 2, pp. 461–464, 1978.

[15] 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 *Proc. IEEE-RAS/RSJ Int. Conf. Humanoid Robots (Humanoids)*, 2007, pp. 557–562.

[16] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. R. Statist. Soc. Series B (Methodological)*, vol. 39, no. 1, pp. 1–38, 1977.

[17] L. R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proc. IEEE*, vol. 77, no. 2, pp. 257–285, Feb. 1989.

[18] K. Fukunaga and D. Olsen, "An algorithm for finding intrinsic dimensionality of data," *IEEE Trans. Comput.*, vol. 20, no. 2, pp. 176–183, 1971.

[19] Y. Kuniyoshi, Y. Ohmura, K. Terada, A. Nagakubo, S. Eitoku, and T. Yamamoto, "Embodied basis of invariant features in execution and perception of whole body dynamic actions—Knacks and focuses of roll-and-rise motion," *Robot. Autonom. Syst.*, vol. 48, no. 4, pp. 189–201, 2004.

Shuhei Ikemoto, Department of Multimedia Engineering, Osaka University, Japan. E-mail: ikemoto@ist.osaka-u.ac.jp.

Heni Ben Amor, Intelligent Autonomous Systems Group, Technische Universitaet Darmstadt, Germany. E-mail: amor@ias.tu-darmstadt.de.

Takashi Minato, ATR Hiroshi Ishiguro Laboratory, Kyoto, Japan. E-mail: minato@atr.jp.

Bernhard Jung, Virtual Reality and Multimedia Group, Technische Universität Bergakademie Freiberg, Germany. E-mail: jung@informatik.tu-freiberg.de.

Hiroshi Ishiguro, Department of Systems Innovation, Osaka University, Japan. E-mail: ishiguro@is.sys.es.osaka-u.ac.jp.

Close physical interaction between robots and humans is a particularly challenging aspect of robot development.

Until recently, robotic systems mostly remained in the realm of industrial applications and academic research.

In recent years, robotics technology significantly matured and produced highly realistic android robots.

The human counterpart is part of the learning system and overall dynamics.

The joints have low mechanical impedance because of the compressibility of air.

The joints have low mechanical impedance because of the compressibility of air.